

**Interval analysis
for
guaranteed nonlinear parameter estimation**

**Eric Walter and Michel Kieffer
{walter, kieffer}@lss.supelec.fr**

27 August 2003

Context

Vector \mathbf{p} of parameters to be estimated from experimental data

Nonlinear estimation

- Classical approach is by minimization of a cost function
- Explicit solution only available in special cases

No explicit solution for most knowledge-based models

Iterative local optimization \rightarrow no guarantee as to results

Need for guaranteed alternatives

Purpose of talk

Show how *interval analysis* allows guaranteed results to be obtained

Numerical solutions provided as
sets guaranteed to contain *all* actual solutions

→ considerable advantage over usual numerical methods

We shall talk of

- Basic concepts of interval analysis
- Tools for nonlinear parameter estimation

Pointers to freely downloadable software are in the proceedings

Basic concepts of interval analysis

Interval

$$[x] = \{x \in \mathbb{R} \mid \underline{x} \leq x \leq \bar{x}\}$$

Width

$$w([x]) = \bar{x} - \underline{x}$$

Midpoint

$$\text{mid}([x]) = \frac{\underline{x} + \bar{x}}{2}$$

Intervals have a dual nature:

- *sets* \Rightarrow set-theoretic operations apply
- *pairs of real numbers* \Rightarrow an arithmetic can be built

Operations on intervals

$$[x] + [y] = [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$$

$$[x] - [y] = [\underline{x} - \bar{y}, \bar{x} - \underline{y}]$$

$$[x] \cdot [y] = [\min\{\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\}, \max\{\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\}]$$

If $0 \notin [y]$ then

$$[x] / [y] = [x] \cdot [1/\bar{y}, 1/\underline{y}]$$

(Specific formulas available for division by interval containing zero)

$[x]$ and $[y]$ treated as if independent variables x and y ,

a major source of pessimism

For instance, since

$$[x] - [y] = [\underline{x} - \bar{y}, \bar{x} - \underline{y}]$$

$[x] - [x]$ not equal to $[0, 0]$, unless $[x]$ degenerate

Multiplication no longer distributive with respect to addition. Instead

$$[x] \cdot ([y] + [z]) \subset [x] \cdot [y] + [x] \cdot [z]$$

known as *subdistributivity* \implies *factorize as much as possible*

Interval counterpart $[f]^*$ of f from \mathbb{R} to \mathbb{R} satisfies

$$[f]^*([x]) = [\{f(x) \mid x \in [x]\}]$$

For any continuous function, $[f]^*([x])$ is the image set $f([x])$

Elementary interval functions expressed in terms of bounds

For instance

$$[\exp]^*([x]) = [\exp(\underline{x}), \exp(\bar{x})]$$

Specific algorithms for

- trigonometric functions
- hyperbolic functions

Interval vector (or *box*) is a Cartesian product of intervals

$$[\mathbf{x}] = [x_1] \times [x_2] \times \cdots \times [x_n] \text{ or } [\mathbf{x}] = ([x_1], [x_2], \dots, [x_n])^T$$

= axis-aligned parallelepiped

Lower bound

$$\underline{\mathbf{x}} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)^T$$

Upper bound

$$\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$$

Width

$$w([\mathbf{x}]) = \max_{1 \leq i \leq n} w([x_i])$$

Midpoint

$$\text{mid}([\mathbf{x}]) = (\text{mid}([x_1]), \dots, \text{mid}([x_n]))^T$$

Classical operations on vectors trivially extend to *interval vectors*

$$\alpha[\mathbf{x}] = (\alpha[x_1]) \times \cdots \times (\alpha[x_n])$$

$$[\mathbf{x}]^T \cdot [\mathbf{y}] = [x_1] \cdot [y_1] + \cdots + [x_n] \cdot [y_n]$$

$$[\mathbf{x}] + [\mathbf{y}] = ([x_1] + [y_1]) \times \cdots \times ([x_n] + [y_n])$$

Interval matrices similarly defined

For example

$$[\mathbf{A}] \cdot [\mathbf{B}] = \left(\sum_{k=1}^n [a_{ik}] \cdot [b_{kj}] \right)_{1 \leq i \leq n, 1 \leq j \leq n}$$

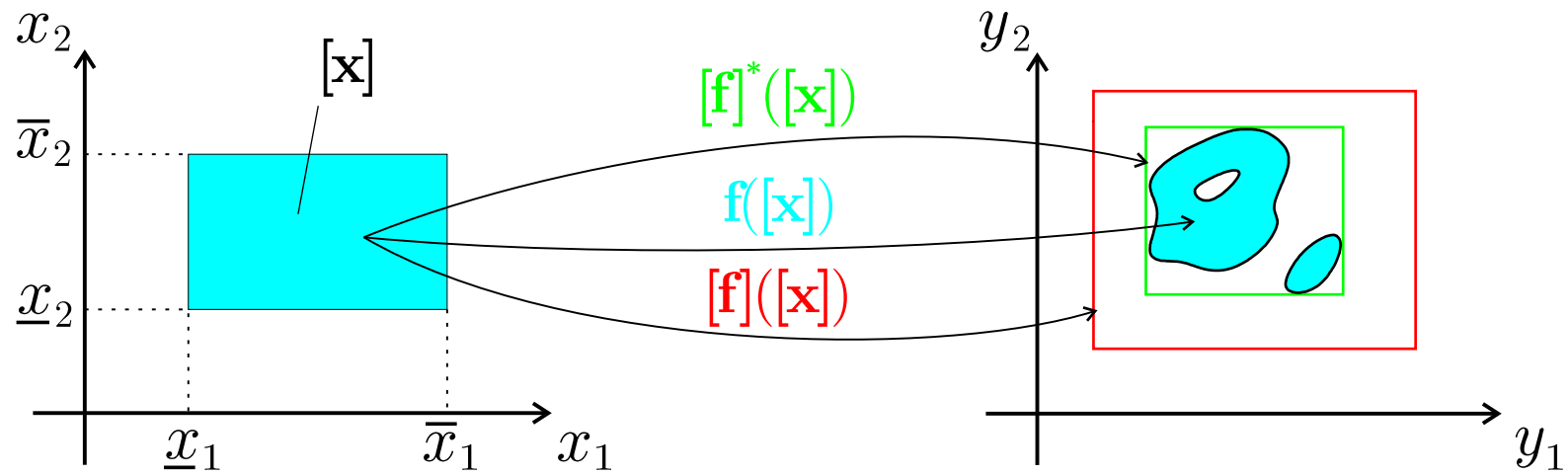
Inclusion functions

$[\mathbf{f}]$ an *inclusion function* for \mathbf{f} if

$$\forall [\mathbf{x}] \in \mathbb{IR}^n, \quad \mathbf{f}([\mathbf{x}]) \subset [\mathbf{f}]([\mathbf{x}])$$

\mathbf{f} may be defined by an algorithm or even by a differential equation

Infinitely many inclusion functions for the same function



$[\mathbf{f}]$ is

convergent if, for any sequence of boxes $[\mathbf{x}]_k$,

$$\lim_{k \rightarrow \infty} w([\mathbf{x}]_k) = 0 \Rightarrow \lim_{k \rightarrow \infty} w([\mathbf{f}]([\mathbf{x}]_k)) = 0$$

minimal if $[\mathbf{f}]([\mathbf{x}])$ is the smallest box that contains $\mathbf{f}([\mathbf{x}])$

inclusion monotonic if

$$[\mathbf{x}] \subset [\mathbf{y}] \Rightarrow [\mathbf{f}]([\mathbf{x}]) \subset [\mathbf{f}]([\mathbf{y}])$$

Minimal inclusion function is inclusion monotonic but not necessarily convergent

Convergent inclusion function may not be inclusion monotonic

Construction of inclusion functions for \mathbf{f}

cast into that of inclusion functions for coordinate functions f_i

\Rightarrow only inclusion functions for $f : \mathbb{R}^n \rightarrow \mathbb{R}$ need be considered

First idea that comes to mind = compute *infimum* and *supremum* of f over box $[\mathbf{x}]$ of interest

\Rightarrow two global optimizations, **usually intractable**

Assume f expressed as composition of

- operators $+$, $-$, \cdot , $/$
- elementary functions \sin , \cos , \exp , sqrt ...

$[f]$ obtained by replacing

- each x_i by $[x_i]$
- each operator or elementary function by interval counterpart

is the *natural inclusion function* of f

(inclusion monotonic)

Example

Four formal expressions of the **same function**

$$f_1(x) = x(x + 1), \quad f_3(x) = x^2 + x,$$

$$f_2(x) = x \cdot x + x, \quad f_4(x) = \left(x + \frac{1}{2}\right)^2 - \frac{1}{4}.$$

On $[x] = [-1, 1]$,

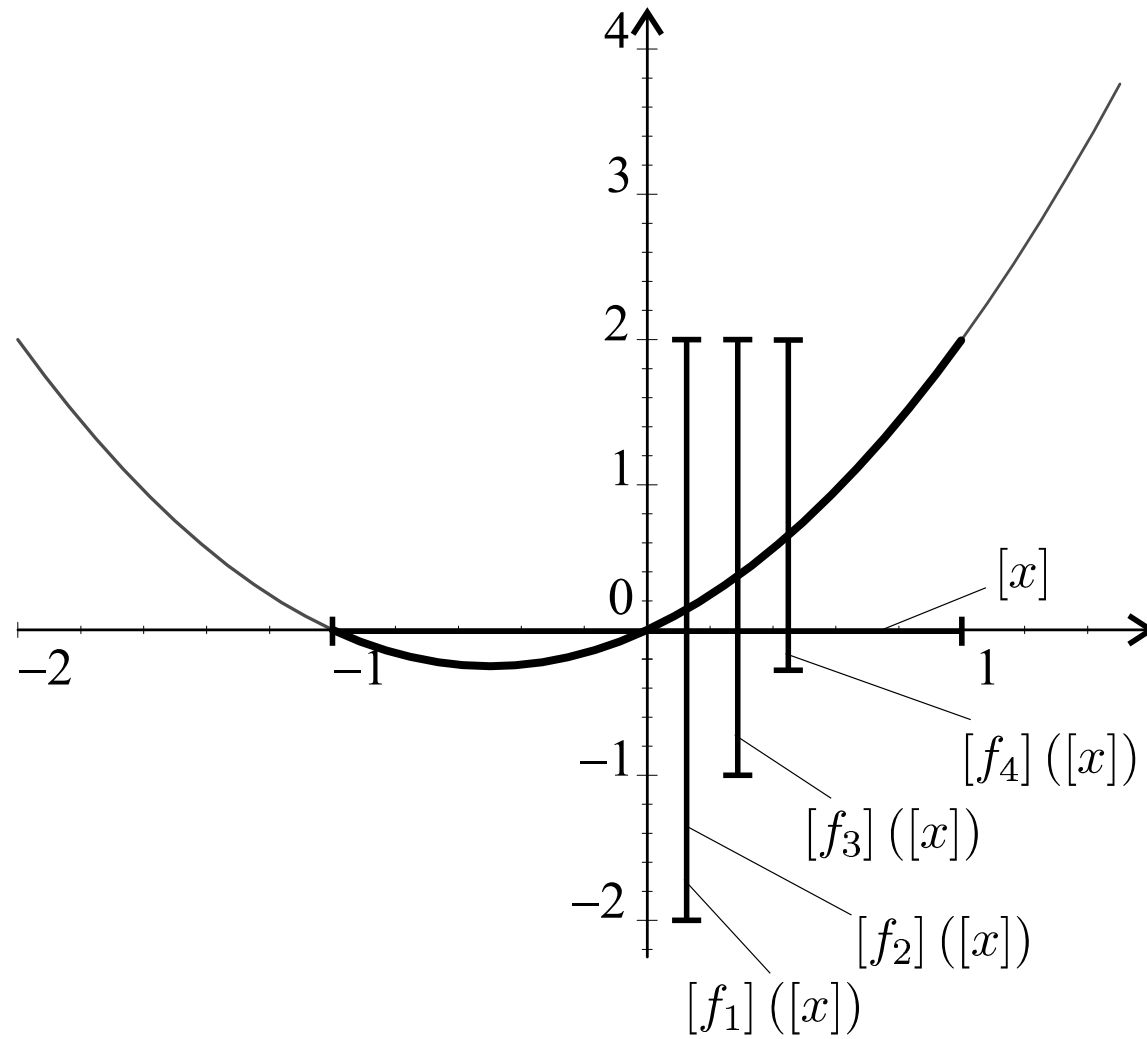
$$[f_1]([x]) = [x]([x] + 1) = [-2, 2],$$

$$[f_2]([x]) = [x] \cdot [x] + [x] = [-2, 2],$$

$$[f_3]([x]) = [x]^2 + [x] = [-1, 2],$$

$$[f_4]([x]) = \left([x] + \frac{1}{2}\right)^2 - \frac{1}{4} = \left[-\frac{1}{4}, 2\right].$$

Only $[f_4](\cdot)$ is minimal $\iff x$ occurs only once



Many other types of inclusion function

If f differentiable over $[\mathbf{x}]$, **mean-value theorem** implies

$$\forall \mathbf{x} \in [\mathbf{x}], \exists \mathbf{z} \in [\mathbf{x}] \text{ such that } f(\mathbf{x}) = f(\mathbf{m}) + \mathbf{g}^T(\mathbf{z}) \cdot (\mathbf{x} - \mathbf{m})$$

with \mathbf{g} the gradient of f and \mathbf{m} the midpoint of $[\mathbf{x}]$

Thus,

$$\forall \mathbf{x} \in [\mathbf{x}], f(\mathbf{x}) \in f(\mathbf{m}) + [\mathbf{g}^T]([\mathbf{x}]) \cdot (\mathbf{x} - \mathbf{m})$$

so

$$f([\mathbf{x}]) \subset f(\mathbf{m}) + [\mathbf{g}^T]([\mathbf{x}]) \cdot ([\mathbf{x}] - \mathbf{m})$$

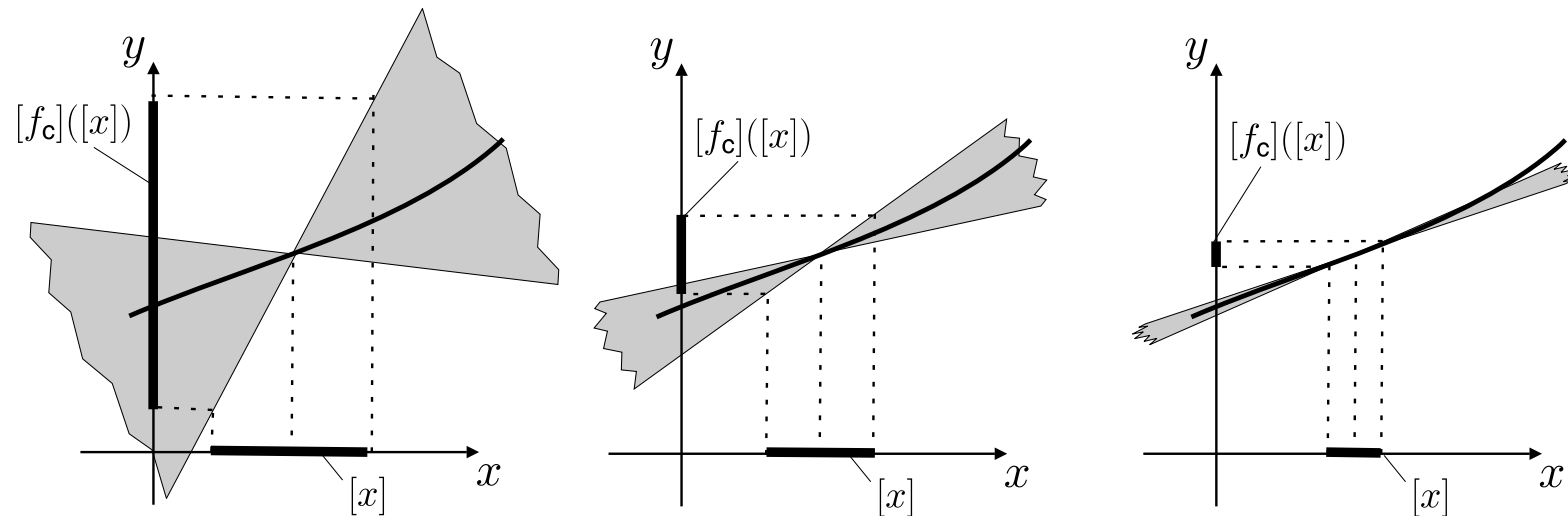
Yields the *centered inclusion function* for f

$$[f_c]([\mathbf{x}]) = f(\mathbf{m}) + [\mathbf{g}^T]([\mathbf{x}]) \cdot ([\mathbf{x}] - \mathbf{m})$$

Centered inclusion function

$$[f_c]([\mathbf{x}]) = f(\mathbf{m}) + [\mathbf{g}^T]([\mathbf{x}]) \cdot ([\mathbf{x}] - \mathbf{m})$$

especially interesting when width of $[\mathbf{x}]$ is small:



For

$$f(x) = x^2 \exp(x) - x \exp(x^2)$$

compare

$[x]$	$f([x])$	$[f]([x])$	$[f]_c([x])$
[0.5, 1.5]	[-4.148, 0]	[-13.82, 9.44]	[-25.07, 25.07]
[0.9, 1.1]	[-0.05380, 0]	[-1.697, 1.612]	[-0.5050, 0.5050]
[0.99, 1.01]	[-0.0004192, 0]	[-0.1636, 0.1628]	[-0.004656, 0.004656]

Subpavings

Intervals and boxes not general enough
to describe all sets \mathbb{S} of interest

Motivates the introduction of subpavings

Subpaving of $[\mathbf{x}]$ = union of nonoverlapping subboxes of $[\mathbf{x}]$

If subpavings $\underline{\mathbb{S}}$ and $\overline{\mathbb{S}}$ such that

$$\underline{\mathbb{S}} \subset \mathbb{S} \subset \overline{\mathbb{S}}$$

then \mathbb{S} bracketed between inner and outer approximations

Distance between $\underline{\mathbb{S}}$ and $\overline{\mathbb{S}}$ indicative of quality of approximation of \mathbb{S}

Computation on subpavings

- allows approximate computation on compact sets
- basic ingredient of estimation algorithms to be presented

Contractors

Consider a vector \mathbf{x} of variables linked by relations (or constraints)

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \quad (1)$$

Assume prior domain for \mathbf{x} is

$$[\mathbf{x}] = [x_1] \times \cdots \times [x_n]$$

Solving (1) for \mathbf{x} in $[\mathbf{x}]$ is a *constraint satisfaction problem* (CSP)

Solution of CSP is

$$\mathbb{S} = \{\mathbf{x} \in [\mathbf{x}] \mid \mathbf{f}(\mathbf{x}) = \mathbf{0}\}$$

Inequality constraints dealt with via *slack variables*

Looking for \mathbb{S} is an NP-complete problem

Instead, *contractors* reduce size of prior domain without losing solutions

Interval Newton contractor classically used for global optimization:

If \mathbf{f} once differentiable, mean-value theorem implies

$$\forall \mathbf{x} \in [\mathbf{x}], \exists \mathbf{z} \in [\mathbf{x}] \mid \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{m}) + \mathbf{J}_{\mathbf{f}}(\mathbf{z}) \cdot (\mathbf{x} - \mathbf{m}) \quad (2)$$

with $\mathbf{J}_{\mathbf{f}}$ the Jacobian matrix of \mathbf{f} and \mathbf{m} the midpoint of $[\mathbf{x}]$.

Assume

- $\hat{\mathbf{x}} \in [\mathbf{x}]$ a solution, so $\mathbf{f}(\hat{\mathbf{x}}) = \mathbf{0}$
- $\#$ constraints equal to $\#$ variables, so $\mathbf{J}_{\mathbf{f}}$ square
- $\mathbf{J}_{\mathbf{f}}$ invertible

then (2) implies

$$\hat{\mathbf{x}} = \mathbf{m} - \mathbf{J}_{\mathbf{f}}^{-1}(\mathbf{z}) \cdot \mathbf{f}(\mathbf{m})$$

Now

$$\hat{\mathbf{x}} = \mathbf{m} - \mathbf{J}_f^{-1}(\mathbf{z}) \cdot \mathbf{f}(\mathbf{m})$$

implies

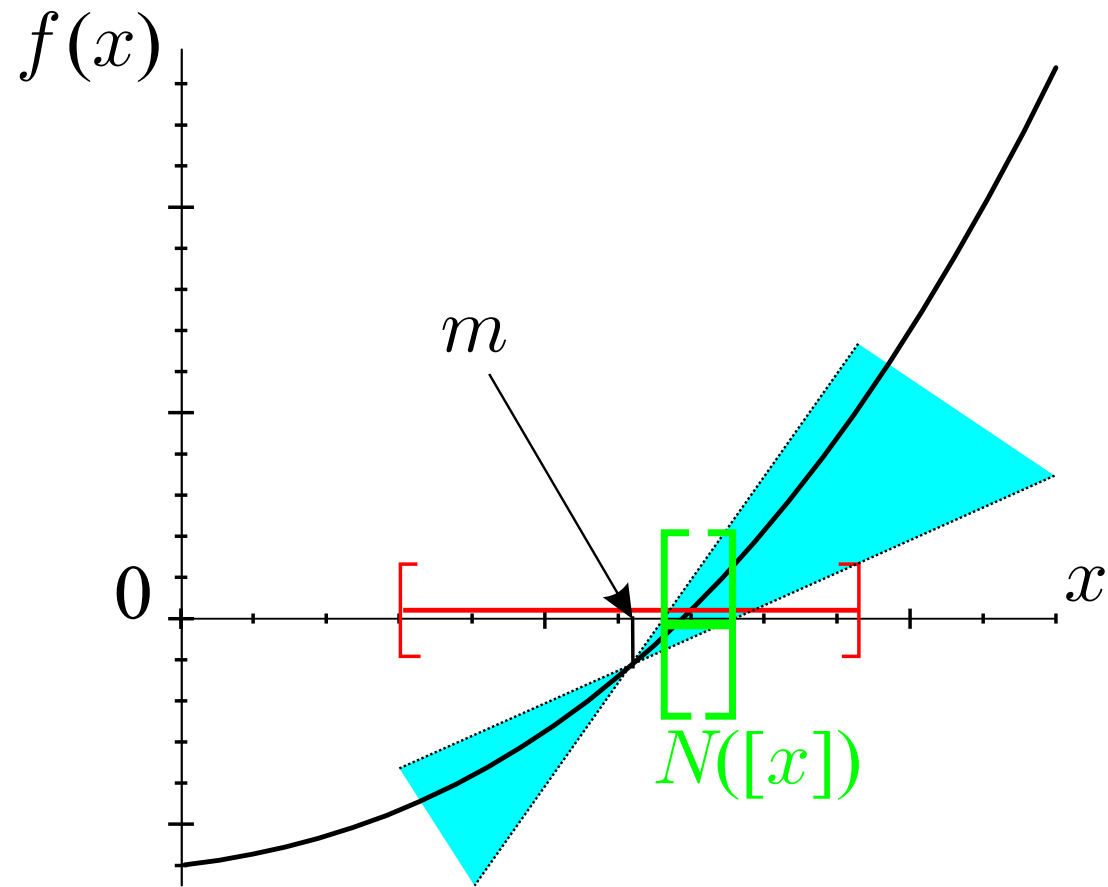
$$\hat{\mathbf{x}} \in \mathbf{m} - \mathbf{J}_f^{-1}([\mathbf{x}]) \cdot \mathbf{f}(\mathbf{m}) \equiv \mathbf{N}([\mathbf{x}])$$

Since $\hat{\mathbf{x}}$ also assumed to belong to $[\mathbf{x}]$, it must belong to

$$[\mathbf{x}_r] = [\mathbf{x}] \cap \mathbf{N}([\mathbf{x}])$$

Interval Newton contractor thus replaces $[\mathbf{x}]$ by $[\mathbf{x}_r]$

$[\mathbf{x}_r]$ may be much smaller (or even empty)



Contractors work without bisection \Rightarrow *escape the curse of dimensionality*

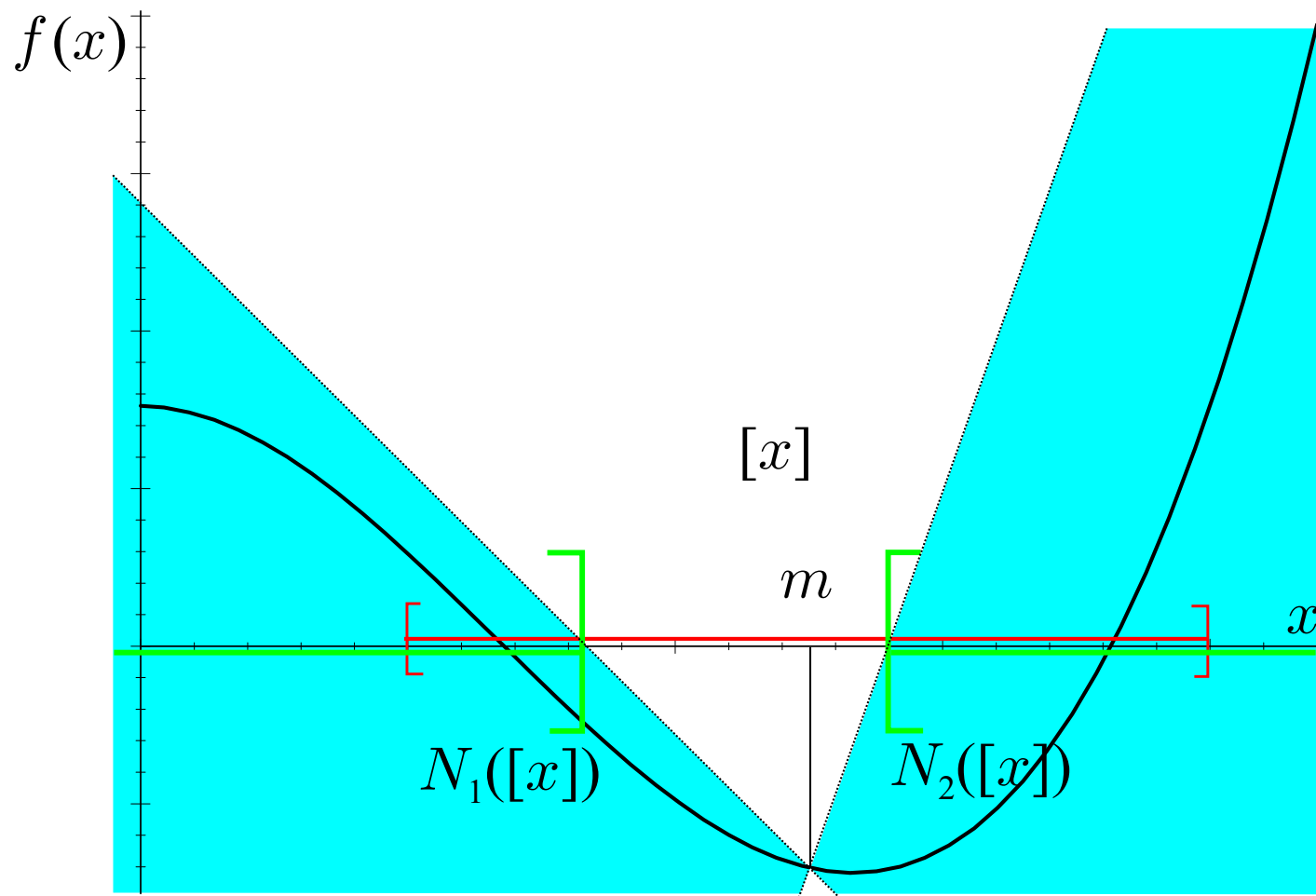
Assumption that $\mathbf{J}_f([\mathbf{x}])$ is invertible can be dropped:

Compute outer approximation of set of all solutions for $\hat{\mathbf{x}}$ of

$$\mathbf{f}(\mathbf{m}) + \mathbf{J}_f([\mathbf{x}]) \cdot (\hat{\mathbf{x}} - \mathbf{m}) = \mathbf{0}$$

a *linear* system of equations

Specific methods involving *preconditioning*



\mathbf{J}_f is not invertible

Basic tools for nonlinear estimation

Global optimization

Interval analysis also used for

- constrained global optimization
- minimax optimization

Here *unconstrained minimization*

Cost function c to be **minimized** over compact set $\mathbb{P} \subset \mathbb{R}^{n_{\mathbb{P}}}$:

$$\min_{\mathbf{p} \in \mathbb{P}} c(\mathbf{p})$$

\hat{c} *global minimum*

\mathbb{S} set of all corresponding *global minimizers*

\mathbb{P} a (possibly very large) domain of interest

Minimizers not expected to lie on its boundary

Main ideas of **Hansen's algorithm** will be presented

Search domain \mathbb{P} assumed to be a box $[\mathbf{p}]_0$ in parameter space

\mathcal{L} = list of boxes in \mathbb{P} that may contain global minimizers

Initially, \mathcal{L} contains only $[\mathbf{p}]_0$

While \mathcal{L} contains at least one box with width greater than ε , do {

1. Select one such box $[\mathbf{p}]$ out of \mathcal{L}
2. Update upper bound $\text{ub}(\hat{c})$ of global minimum \hat{c}
3. Try to eliminate $[\mathbf{p}]$ via midpoint, monotonicity and convexity tests
4. Should this fails, try to eliminate or at least reduce $[\mathbf{p}]$ with interval Newton contractor
5. If $[\mathbf{p}]$ resists elimination and $w([\mathbf{p}]) > \varepsilon$ then bisect it and push resulting boxes into \mathcal{L} . Otherwise push $[\mathbf{p}]$ back as is into \mathcal{L} }

At last eliminate all boxes of \mathcal{L} that no longer pass midpoint test

Step 1

Strategy for **box selection** :

take $[\mathbf{p}]$ with the largest potential,

i.e., such that

$$\forall [\mathbf{q}] \text{ in } \mathcal{L} \quad \text{lb}([c]([\mathbf{p}])) \leq \text{lb}([c]([\mathbf{q}])).$$

Step 2

Update upper bound of \hat{c} by local search

Step 3

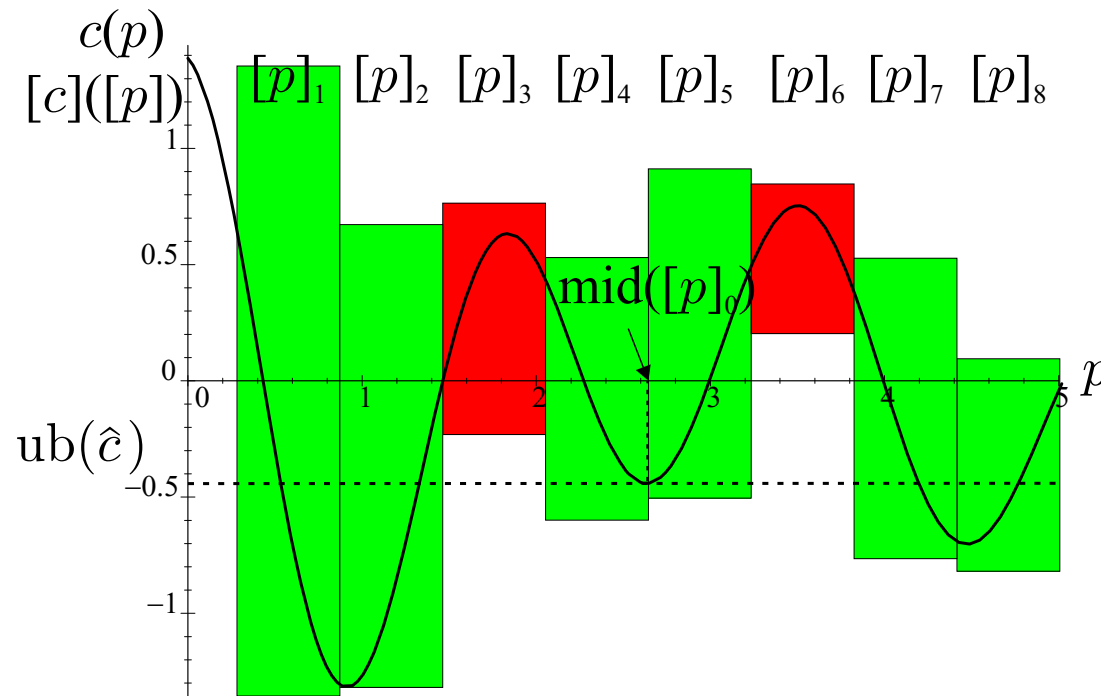
Three opportunities to **eliminate $[\mathbf{p}]$** at Step 3

Midpoint test

If

$$\text{lb}([c]([p])) > \text{ub}(\hat{c}),$$

then no \mathbf{p} in $[p]$ can be a global minimizer so $[p]$ can be eliminated

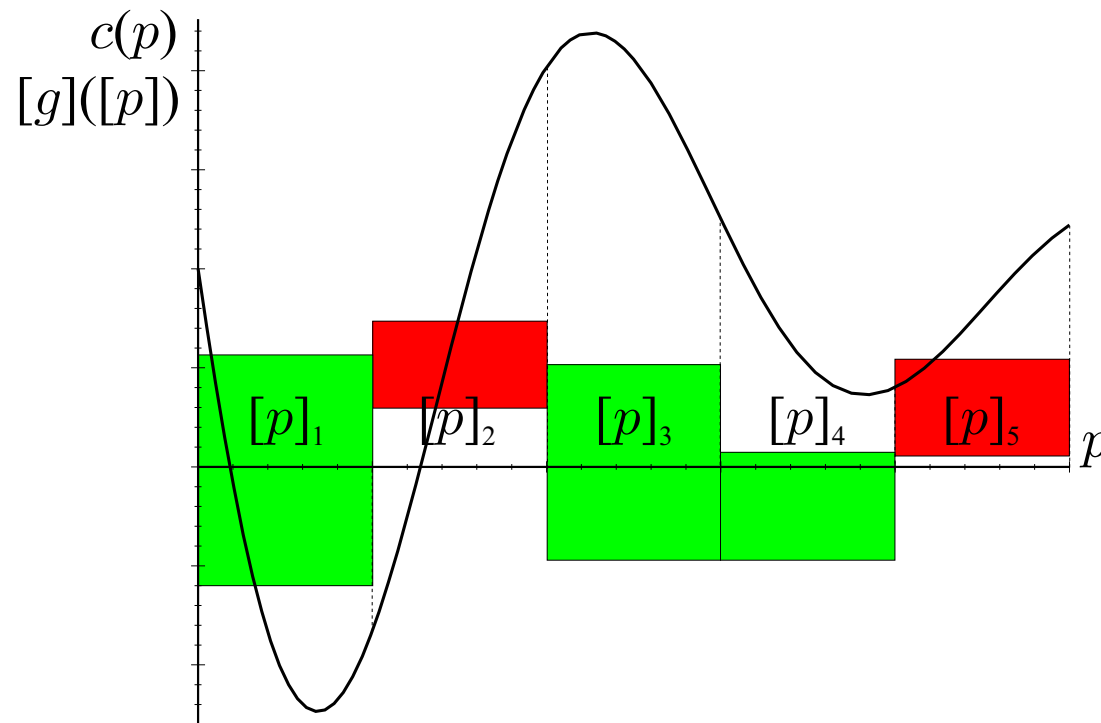


Monotonicity test

If

$$\mathbf{0} \notin [g]([p]),$$

then no \mathbf{p} in $[p]$ can be an unconstrained minimizer, so $[p]$ can be eliminated



Convexity test

Based on nonnegativity of Hessian \mathbf{H} at any minimizer

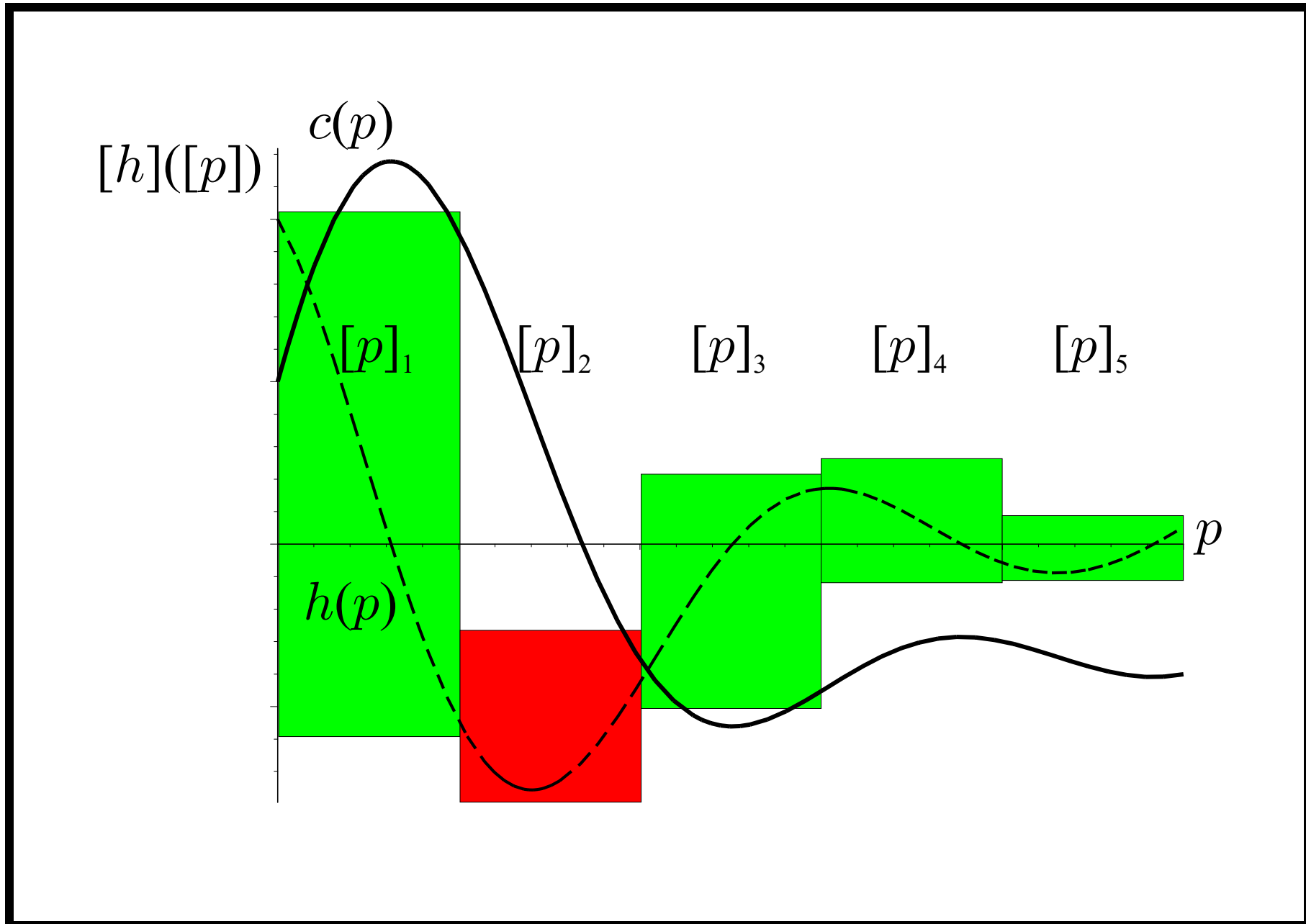
Necessary condition: all diagonal entries of \mathbf{H} satisfy

$$h_{ii} \geq 0, i = 1, \dots, n_p.$$

So if

$$\exists i \in \{1, \dots, n_p\} \mid [h_{ii}]([\mathbf{p}]) < 0,$$

then no \mathbf{p} in $[\mathbf{p}]$ can be an unconstrained minimizer and $[\mathbf{p}]$ can be eliminated



- *Automatic differentiation* used to compute inclusion functions for \mathbf{g} and \mathbf{H} from code evaluating $c(\mathbf{p})$
- If c not differentiable, or if inclusion functions for \mathbf{g} or \mathbf{H} not available, parts of Step 3 dropped
- In the limit, Step 3 may be based on midpoint test only

Step 4 uses **interval Newton contractor** to reduce size of $[\mathbf{p}]$ without bisection.

Problem is to find a box guaranteed to contain all solutions in $[\mathbf{p}]$ of

$$\mathbf{g}(\mathbf{p}) = \mathbf{0}$$

so

- \mathbf{g} stands for \mathbf{f}
- \mathbf{H} stands for \mathbf{J}_f

Possible only if c twice differentiable

- Since $\varepsilon > 0$, Hansen's algorithm stops after a **finite number of steps**
- Set \mathbb{S} of **all global minimizers** of c **included in the union of the boxes in \mathcal{L}**
- **Possible to compute an interval $[\hat{c}]$ guaranteed to contain \hat{c}**
 - lower bound

$$\underline{\hat{c}} = \min_{[\mathbf{p}] \in \mathcal{L}} lb([c]([\mathbf{p}]))$$

- upper bound the best upper bound obtained at Step 2

Set inversion

Let

- \mathbf{f} be a possibly nonlinear function from \mathbb{R}^{n_p} to \mathbb{R}^{n_y}
- \mathbb{Y} be a subpaving of \mathbb{R}^{n_y}

Set inversion is the characterization of the reciprocal image of \mathbb{Y}

$$\mathbb{S} = \{\mathbf{p} \in \mathbb{R}^{n_p} \mid \mathbf{f}(\mathbf{p}) \in \mathbb{Y}\} = \mathbf{f}^{-1}(\mathbb{Y})$$

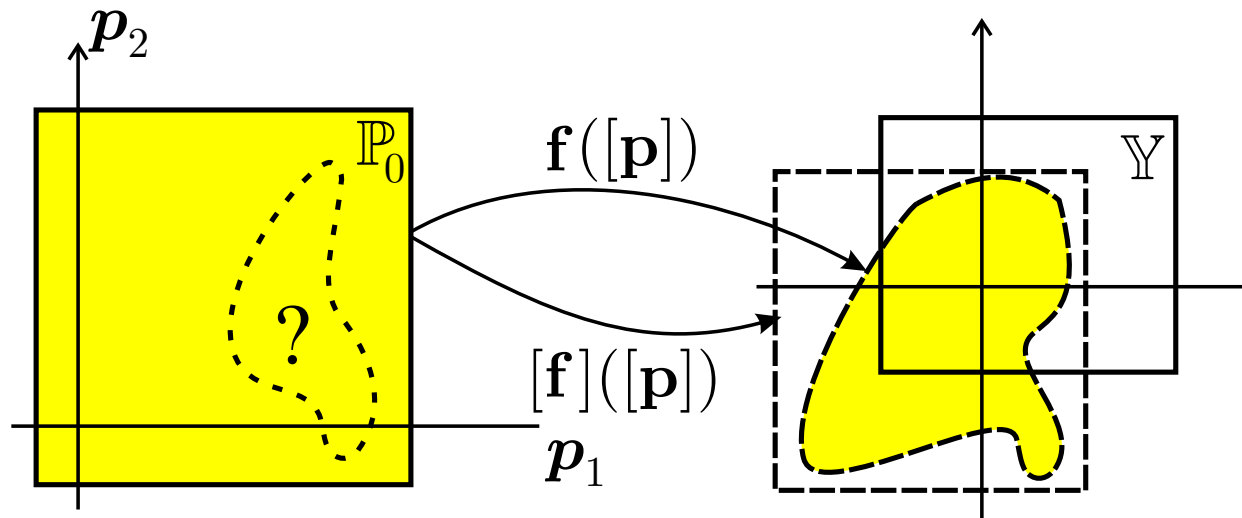
Using

- an inclusion function $[\mathbf{f}]$ for \mathbf{f} ,
- a (possibly very large) search box $[\mathbf{p}]_0$

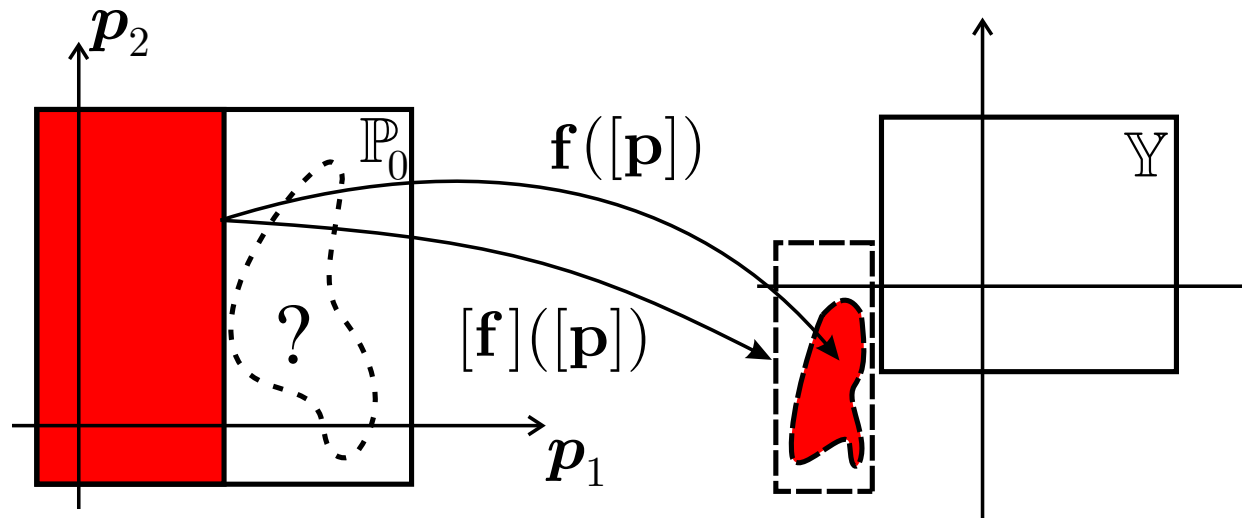
SIVIA, for *Set Inverter Via Interval Analysis*, computes subpavings $\underline{\mathbb{S}}$ and $\overline{\mathbb{S}}$ such that

$$\underline{\mathbb{S}} \subset \mathbb{S} \subset \overline{\mathbb{S}}$$

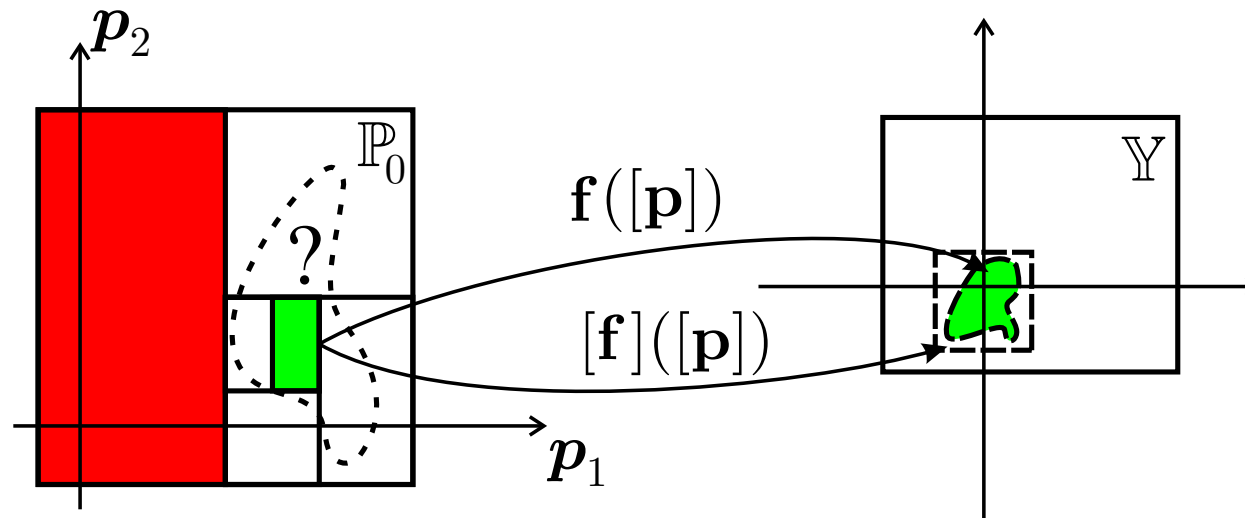
by successive bisections and selections



Yellow box is *undetermined*



Red box **proven** to be outside \mathcal{S}



Green box **proven** to be inside \mathcal{S}

Algorithm SIVIA(in: $\mathbf{f}, \mathbb{Y}, [\mathbf{p}]$, ε ; inout: $\underline{\mathbb{S}}, \bar{\mathbb{S}}$)

```

1  if  $[\mathbf{f}]([\mathbf{p}]) \cap \mathbb{Y} = \emptyset$  return;
2  if  $[\mathbf{f}]([\mathbf{p}]) \subset \mathbb{Y}$  then
3  { $\underline{\mathbb{S}} := \underline{\mathbb{S}} \cup [\mathbf{p}]$ ;  $\bar{\mathbb{S}} := \bar{\mathbb{S}} \cup [\mathbf{p}]$ ; return;};
4  if  $w([\mathbf{p}]) < \varepsilon$  then { $\bar{\mathbb{S}} := \bar{\mathbb{S}} \cup [\mathbf{p}]$ ; return;};
5  SIVIA( $\mathbf{f}, \mathbb{Y}, L[\mathbf{p}]$ ,  $\varepsilon, \underline{\mathbb{S}}, \bar{\mathbb{S}}$ );
   SIVIA( $\mathbf{f}, \mathbb{Y}, R[\mathbf{p}]$ ,  $\varepsilon, \underline{\mathbb{S}}, \bar{\mathbb{S}}$ ).

```

All boxes in *uncertainty layer* $\Delta\mathbb{S}$ between $\underline{\mathbb{S}}$ and $\bar{\mathbb{S}}$
 have a width smaller than ε

Guaranteed numerical integration

Assume **model of interest is the ODE**

$$\mathbf{x}' = \mathbf{g}(\mathbf{x}, \mathbf{p}, t), \text{ with } \mathbf{x}(0) = \mathbf{x}_0(\mathbf{p}) \quad (3)$$

where \mathbf{p} only known to belong to $[\mathbf{p}]$

Let $\mathbf{f}(\mathbf{p}, t)$ be the solution of (3) for a given $\mathbf{p} \in [\mathbf{p}]$

Guaranteed integrator computes set containing $\mathbf{f}([\mathbf{p}], t)$

Guaranteed integrators based on interval analysis readily available,
e.g., AWA, COSY or VNODE

Well suited when $[\mathbf{p}]$ a degenerate box with zero width

For large boxes, as needed in the context of parameter estimation,
enclosure for $\mathbf{f}([\mathbf{p}], t)$ may become *very* pessimistic

Solution: bound (if possible) model between *cooperative systems*

The dynamical system

$$\mathbf{x}' = \mathbf{g}(\mathbf{x}, t)$$

where $\mathbf{x} \in \mathcal{D} \subset \mathbb{R}^n$, is *cooperative over \mathcal{D}* if

$$\frac{\partial g_i(\mathbf{x}, t)}{\partial x_j} \geq 0 \text{ for all } i \neq j, t \geq 0 \text{ and } \mathbf{x} \in \mathcal{D}$$

Theorem 1 *If there exists a pair of cooperative systems*

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

satisfying

$$\underline{\mathbf{g}}(\mathbf{x}, \underline{\mathbf{p}}, \bar{\mathbf{p}}, t) \leq \mathbf{g}(\mathbf{x}, \mathbf{p}, t) \leq \bar{\mathbf{g}}(\mathbf{x}, \underline{\mathbf{p}}, \bar{\mathbf{p}}, t)$$

for all $\mathbf{p} \in [\underline{\mathbf{p}}, \bar{\mathbf{p}}]$, $t \geq 0$ and $\mathbf{x} \in \mathcal{D}$, and if

$$\underline{\mathbf{x}}_0(\underline{\mathbf{p}}, \bar{\mathbf{p}}) \leq \mathbf{x}_0(\mathbf{p}) \leq \bar{\mathbf{x}}_0(\underline{\mathbf{p}}, \bar{\mathbf{p}})$$

for all $\mathbf{p} \in [\underline{\mathbf{p}}, \bar{\mathbf{p}}]$, then

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

In theorem, $\underline{\mathbf{x}}(t)$ is the flow $\underline{\varphi}(\underline{\mathbf{p}}, \bar{\mathbf{p}}, t)$ associated with

$$\{\underline{\mathbf{x}}' = \underline{\mathbf{g}}(\underline{\mathbf{x}}, \underline{\mathbf{p}}, \bar{\mathbf{p}}, t), \underline{\mathbf{x}}(0) = \underline{\mathbf{x}}_0(\underline{\mathbf{p}}, \bar{\mathbf{p}})\}$$

and $\bar{\mathbf{x}}(t)$ the flow $\bar{\varphi}(\underline{\mathbf{p}}, \bar{\mathbf{p}}, t)$ associated with

$$\{\bar{\mathbf{x}}' = \bar{\mathbf{g}}(\bar{\mathbf{x}}, \underline{\mathbf{p}}, \bar{\mathbf{p}}, t), \bar{\mathbf{x}}(0) = \bar{\mathbf{x}}_0(\underline{\mathbf{p}}, \bar{\mathbf{p}})\}$$

Box-valued function

$$[\varphi](\underline{\mathbf{p}}, \bar{\mathbf{p}}, t) = [\underline{\varphi}(\underline{\mathbf{p}}, \bar{\mathbf{p}}, t), \bar{\varphi}(\underline{\mathbf{p}}, \bar{\mathbf{p}}, t)]$$

thus an inclusion function for solution of ODE

Application to
nonlinear parameter estimation

Parameter optimization

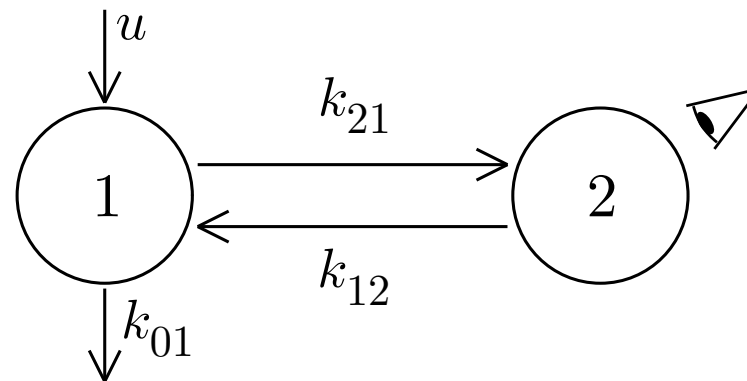
Nonlinear estimation often boils down to minimizing some cost

Cost not convex in general \Rightarrow

- local minimization
- global search based on random exploration

cannot offer any guarantee

Guaranteed interval algorithms particularly attractive



Two-compartment model

State equations readily obtained from conservation law as

$$\mathbf{x}' = \mathbf{g}(\mathbf{x}, \mathbf{p}, u)$$

where

$$\mathbf{p} = (k_{01}, k_{12}, k_{21})^T$$

and

$$\mathbf{g}(\mathbf{x}, \mathbf{p}, u) = \begin{pmatrix} -(k_{21} + k_{01})x_1 + k_{12}x_2 + u \\ k_{21}x_1 - k_{12}x_2 \end{pmatrix}$$

Quantity x_2 of material in Compartment 2 is observed, so

$$y_m(\mathbf{p}, t_i) = x_2(\mathbf{p}, t_i), \quad i = 1, \dots, n_y$$

No input ($u \equiv 0$) and initial condition is $\mathbf{x}_0 = (1, 0)^T$

Then,

$$y_m(\mathbf{p}, t_i) = \alpha(\mathbf{p}) \left(e^{\lambda_1(\mathbf{p})t_i} - e^{\lambda_2(\mathbf{p})t_i} \right)$$

where

$$\alpha(\mathbf{p}) = \frac{k_{21}}{\sqrt{(k_{01} - k_{12} + k_{21})^2 + 4k_{12}k_{21}}}$$

$$\lambda_{1,2}(\mathbf{p}) = -\frac{1}{2} [(k_{01} + k_{12} + k_{21}) \pm ((k_{01} - k_{12} + k_{21})^2 + 4k_{12}k_{21})^{1/2}]$$

Model is linear, but its output depends nonlinearly on \mathbf{p}

Least-square estimate of \mathbf{p} obtained by minimizing

$$c(\mathbf{p}) = \sum_{i=1}^{n_y} (y(t_i) - y_m(\mathbf{p}, t_i))^2$$

- sixteen data points
- $\varepsilon = 10^{-9}$
- about *one day* on a Pentium at 200MHz to enclose all global minimizers in two tiny boxes in parameter space!

- Two boxes because \mathbf{p} *only locally identifiable*
- *Identifiability study bypassed*
- Disappointing length of computations due to multioccurrences of parameters in natural inclusion function for the cost
- *Optimization time drastically decreased to about 90s* via intermediate parametrization of $y_m(\mathbf{p}, t_i)$ in terms of $\alpha(\mathbf{p})$, $\lambda_1(\mathbf{p})$ and $\lambda_2(\mathbf{p})$

Parameter bounding

One now looks for the *set* of all parameter vectors that are consistent with

- experimental data
- model structure
- error bounds

Experimental datum $y(t_i)$ corresponds to a known interval $[\underline{e}_i, \bar{e}_i]$ of acceptable errors

$\mathbf{p} \in [\mathbf{p}]_0$ *acceptable* if

$$\underline{e}_i \leq y(t_i) - y_m(\mathbf{p}, t_i) \leq \bar{e}_i \text{ for all } i = 1, \dots, n_y$$

Parameter estimation amounts to characterizing

$$\begin{aligned}\mathcal{S} &= \{\mathbf{p} \in [\mathbf{p}]_0 \mid \mathbf{p} \text{ is acceptable}\} \\ &= \{\mathbf{p} \in [\mathbf{p}]_0 \mid \mathbf{y}_m(\mathbf{p}) \in [\mathbf{y}]\},\end{aligned}$$

with

$$\begin{aligned}[\mathbf{y}] &= [y(t_1) - \bar{e}_1, y(t_1) - \underline{e}_1] \times \\ &\quad \cdots \times [y(t_{n_y}) - \bar{e}_{n_y}, y(t_{n_y}) - \underline{e}_{n_y}]\end{aligned}$$

and

$$\mathbf{y}_m(\mathbf{p}) = (y_m(\mathbf{p}, t_1), \dots, y_m(\mathbf{p}, t_{n_y}))^T$$

Multioccurrences of \mathbf{p} due to cost function avoided

\Rightarrow pessimism decreased

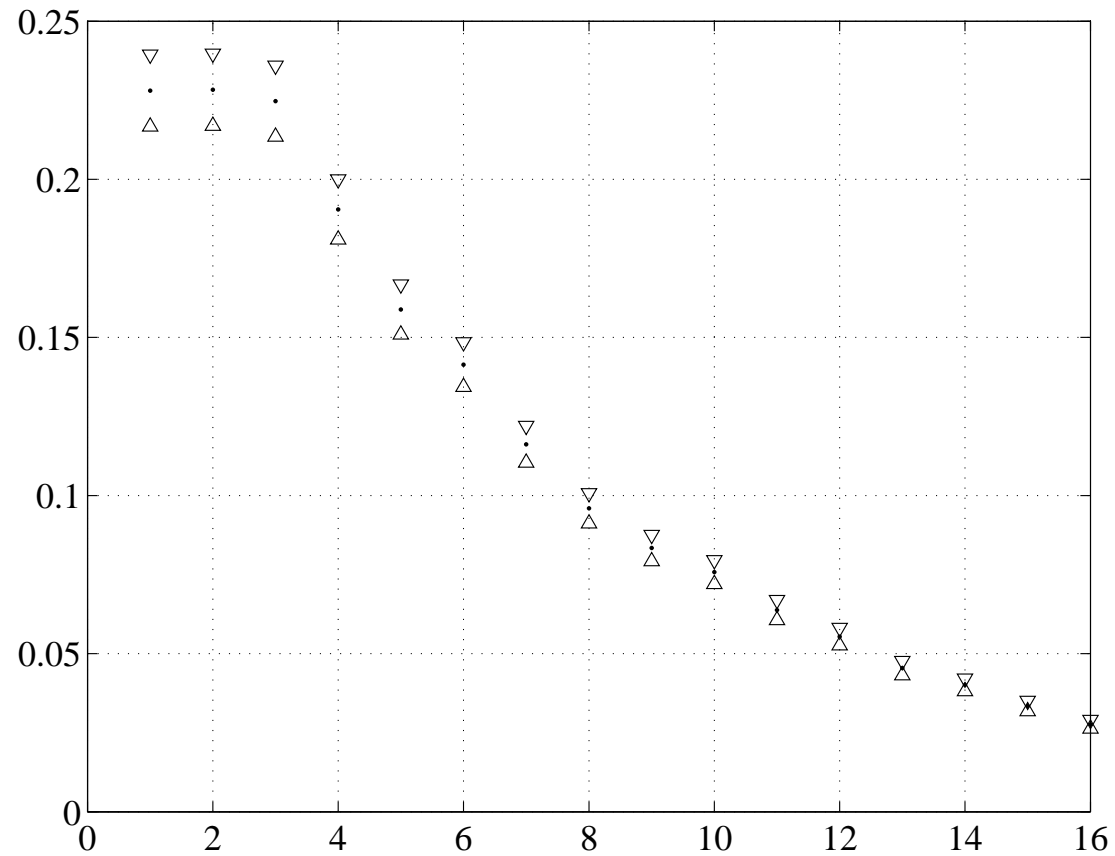
Guaranteed enclosure of \mathcal{S} obtained with SIVIA

Two approaches to be considered

- via a closed-form expression for $y_m(\mathbf{p}, t_i)$
- via guaranteed numerical integration

illustrated on the same compartmental model

Using a closed-form expression



Interval data (true system is linear)

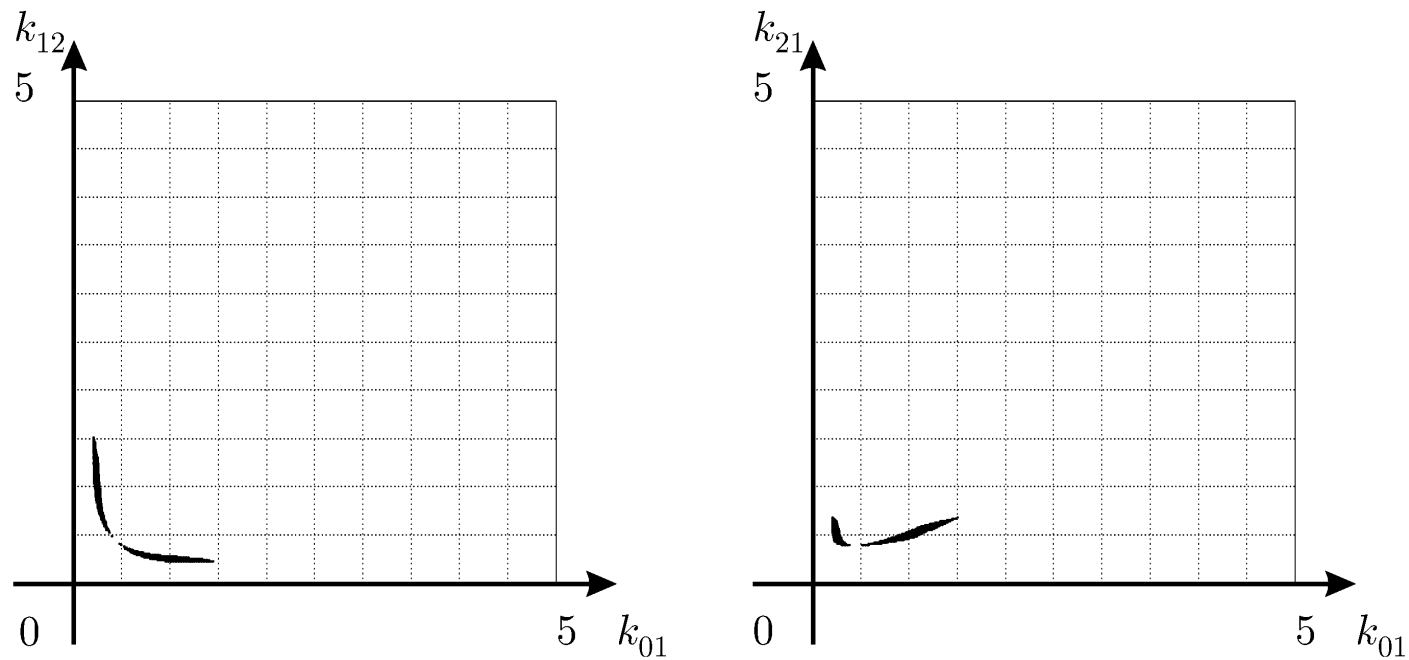
Using SIVIA in $[\mathbf{p}]_0 = [0, 5]^{\times 3}$ leads to

ε	0.005	0.0025	0.00125
Comput. time (s)	9	14	24
Volume of $\bar{\mathcal{S}}$	$1.7 \cdot 10^{-3}$	$4 \cdot 10^{-4}$	$1.2 \cdot 10^{-4}$

Table 1: Results using a closed-form expression

Computations on Athlon 1800+

Projections of $\bar{\mathbb{S}}$ using a closed-form expression with $\varepsilon = 0.0025$



Again a consequence of lack of global identifiability

Using guaranteed integration

Closed-form expression for $y_m(\mathbf{p}, t_i)$ **no longer needed**

For any $[\mathbf{p}] = [\underline{\mathbf{p}}, \bar{\mathbf{p}}]$ such that $\underline{\mathbf{p}} \geq \mathbf{0}$,

$\mathbf{g}(\mathbf{x}, \mathbf{p}, u)$ enclosed between

$$\underline{\mathbf{g}}(\mathbf{x}, \underline{\mathbf{p}}, \bar{\mathbf{p}}, u) = \begin{pmatrix} -(\bar{k}_{21} + \bar{k}_{01})x_1 + \underline{k}_{12}x_2 + u \\ \underline{k}_{21}x_1 - \bar{k}_{12}x_2 \end{pmatrix}$$

and

$$\bar{\mathbf{g}}(\mathbf{x}, \underline{\mathbf{p}}, \bar{\mathbf{p}}, u) = \begin{pmatrix} -(\underline{k}_{21} + \underline{k}_{01})x_1 + \bar{k}_{12}x_2 + u \\ \bar{k}_{21}x_1 - \underline{k}_{12}x_2 \end{pmatrix}$$

Since

$$\mathbf{x}' = \underline{\mathbf{g}}(\mathbf{x}, \underline{\mathbf{p}}, \overline{\mathbf{p}}, u)$$

and

$$\mathbf{x}' = \overline{\mathbf{g}}(\mathbf{x}, \underline{\mathbf{p}}, \overline{\mathbf{p}}, u)$$

cooperative, easy to get an inclusion function for $y_m(\mathbf{p}, t_i)$

Interval data are as before

SIVIA + toolbox VNODE now lead to

ε	0.01	0.005
Comput. time (s)	1300	1600
Volume of $\bar{\mathcal{S}}$	$2.5 \cdot 10^{-3}$	$6 \cdot 10^{-4}$

Table 2: Results with guaranteed integration

Shape and volume for $\varepsilon = 0.005$ similar to those obtained with closed-form solution for $\varepsilon = 0.0025$

For the same accuracy, computing time using guaranteed integration more than 100 times larger than with closed-form expression

Nonlinear system

Assume now that

$$k_{01}(x_1) = \frac{a}{1 + bx_1}$$

(Michaelis-Menten nonlinearity)

State equation becomes nonlinear

$$\mathbf{x}' = \mathbf{h}(\mathbf{x}, \mathbf{p}, u)$$

where

$$\mathbf{p} = (a, b, k_{12}, k_{21})^T$$

and

$$\mathbf{h}(\mathbf{x}, \mathbf{p}, u) = \begin{pmatrix} -k_{21}x_1 - \frac{ax_1}{1 + bx_1} + k_{12}x_2 + u \\ k_{21}x_1 - k_{12}x_2 \end{pmatrix}$$

Again

- compartment 2 is observed, with input and initial conditions as before
- inclusion function based on **guaranteed numerical integration can be employed**

For any $\mathbf{p} \in [\underline{\mathbf{p}}, \overline{\mathbf{p}}]$ such that $\underline{\mathbf{p}} \geq \mathbf{0}$, possible to bound $\mathbf{h}(\mathbf{x}, \mathbf{p}, u)$ between

$$\begin{pmatrix} - \left(\overline{k}_{21} + \frac{\overline{a}}{1 + \underline{b}x_1} \right) x_1 + \underline{k}_{12}x_2 + u \\ \underline{k}_{21}x_1 - \overline{k}_{12}x_2 \end{pmatrix}$$

and

$$\begin{pmatrix} - \left(\underline{k}_{21} + \frac{\underline{a}}{1 + \overline{b}x_1} \right) x_1 + \overline{k}_{12}x_2 + u \\ \overline{k}_{21}x_1 - \underline{k}_{12}x_2 \end{pmatrix}$$

Resulting systems are cooperative, as $\underline{\mathbf{p}} \geq \mathbf{0}$

Inclusion function for $y_m(\mathbf{p}, t_i)$ built by guaranteed numerical integration

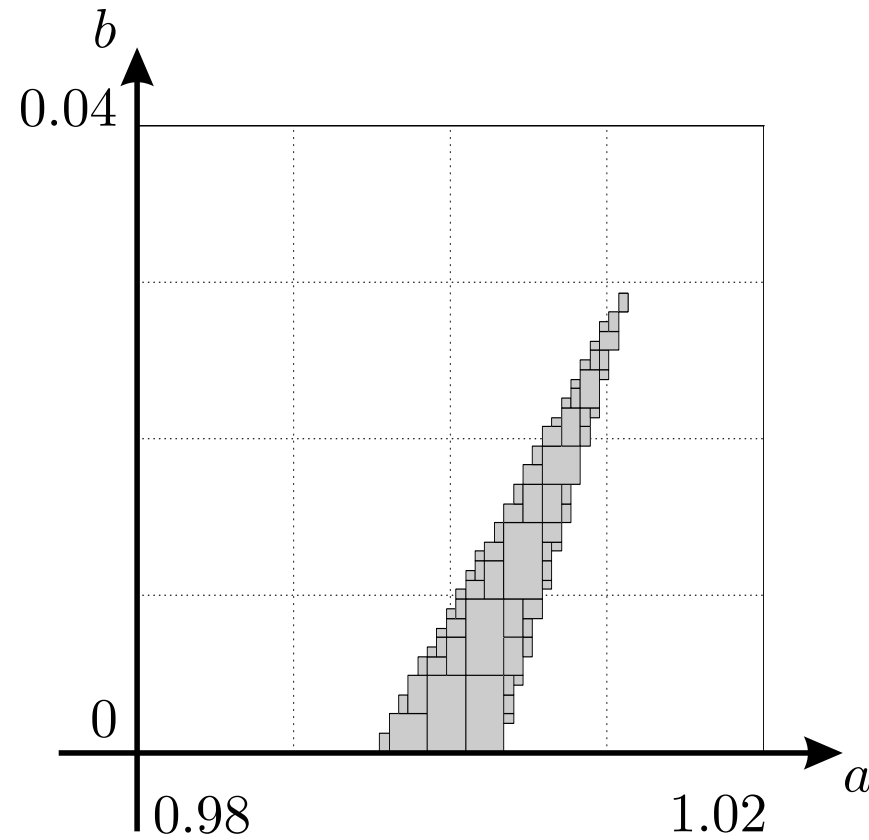
Two sets of data points considered

First data set: data generated by the same *linear* model as before

SIVIA now used with initial search box

$$[\mathbf{p}]_0 = [0, 5] \times [0, 5] \times [0.25, 0.25] \times [0.5, 0.5]$$

so k_{12} and k_{21} treated as known *a priori*



Outer approximation of solution set for (a, b) (true system is linear)

By projection of $\bar{\mathbb{S}}$, one gets

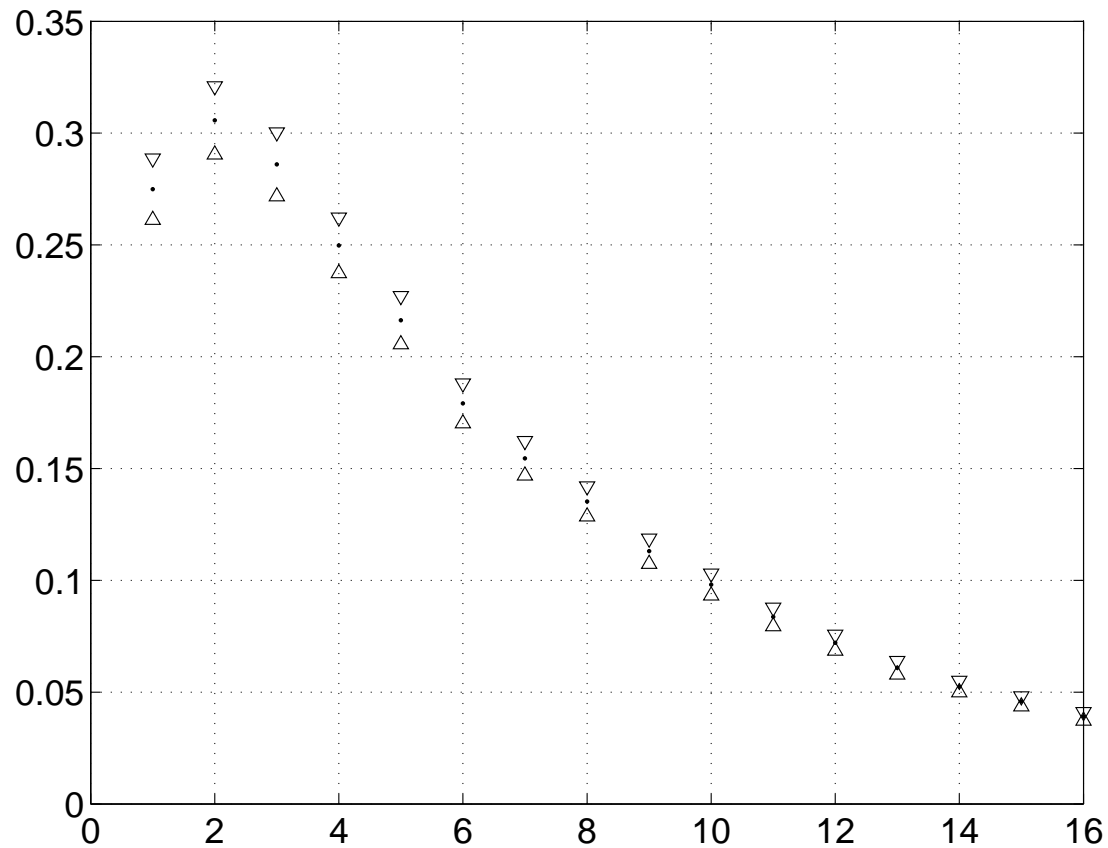
$$[a] = [0.9955, 1.0114]$$

and

$$[b] = [0, 0.02930]$$

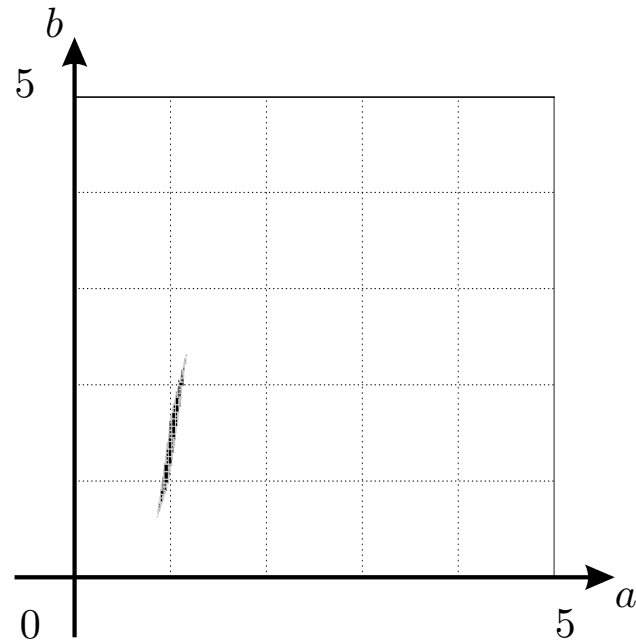
Since data generated with a linear model,
it comes as no surprise that $[b]$ includes 0

Second data set generated by a *nonlinear* system



Interval data (**true system is nonlinear**)

Initial search box as for first data set



Outer approximation of solution set for (a, b)
(true system is nonlinear)

As b cannot be zero, data could not have been generated by a linear model, given hypotheses

Conclusions and perspectives

Global deterministic methods based on interval analysis have definite advantages over more conventional local iterative methods, which are unable to provide guaranteed results

Structural identifiability studies can be bypassed since all solutions are provided

Examples have shown that it is possible to estimate parameters of models defined by (possibly nonlinear) ODEs

Main challenge is increasing complexity of tractable problems

Two allies in this endeavor have been briefly presented

- **contractors** allow boxes to be reduced and sometimes eliminated without bisection
- **cooperativity** allows efficient inclusion functions to be derived for ODEs

The ideas presented here in the context of parameter identification readily extend to state estimation or parameter tracking