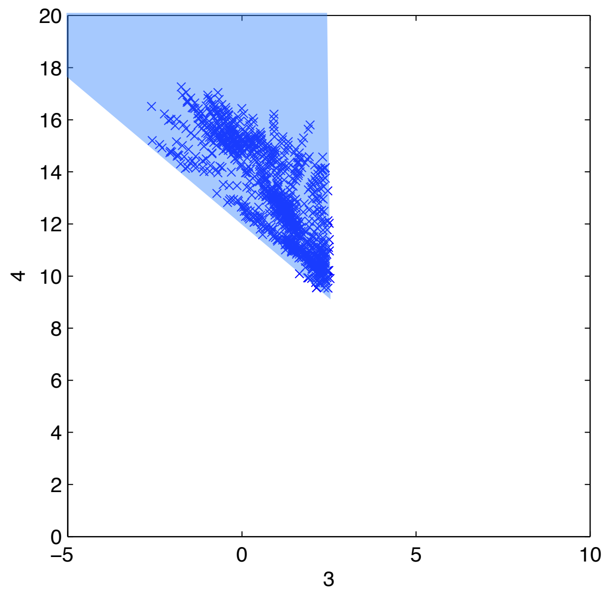
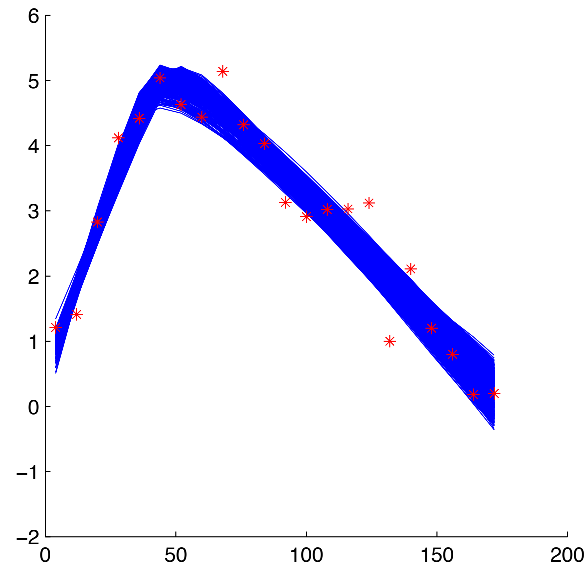
Cluster Newton Method for Sampling Multiple Solutions of an Underdetermined Inverse Problem: Parameter Identification for Pharmacokinetics





Yasunori Aoki
(University of Waterloo)
Uppsala University

Ken Hayami
National Institute of Informatics

Hans De Sterck
University of Waterloo

Akihiko Konagaya
Tokyo Institute of Technology

Research funded and supported by

Ministry of Education, Sports, Science and Technology, Japan

National Institute of Informatics, Tokyo, Japan

Natural Science and Engineering Research Council of Canada

Model Parameter Estimation

Finitely Parameterized Mathematical Model

$$\frac{du_1}{dt} = -x_1 u_1 u_4$$

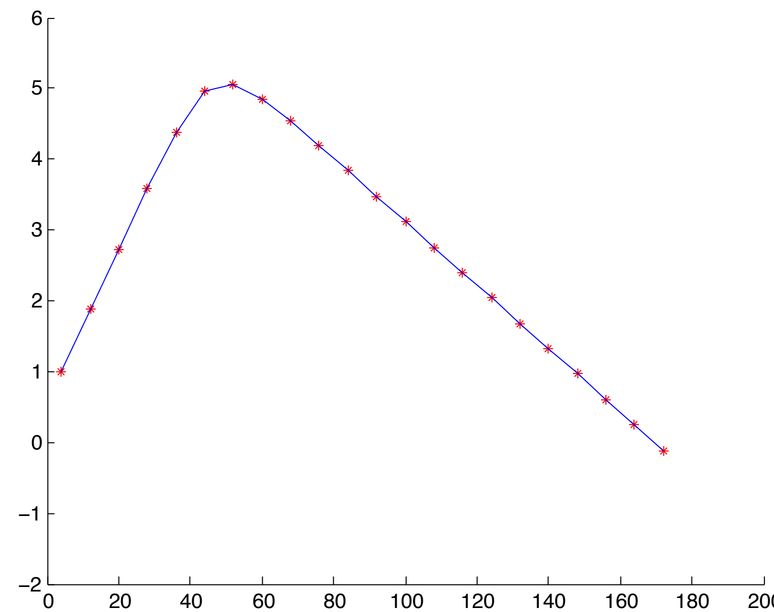
$$\frac{du_2}{dt} = x_1 u_1 u_4 - \frac{1}{x_2} u_2$$

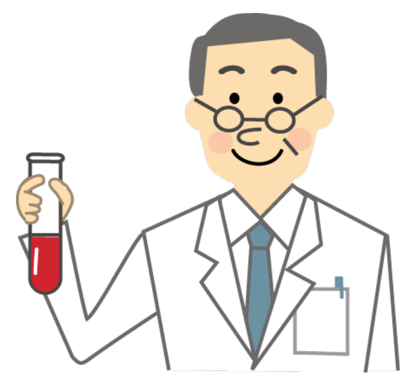
$$\frac{du_3}{dt} = \frac{1}{x_2} u_2 - \frac{1}{x_3} u_3$$

$$\frac{du_4}{dt} = \frac{x_4}{x_5} u_3 - x_6 u_4$$

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}^*$$

Experimental Data





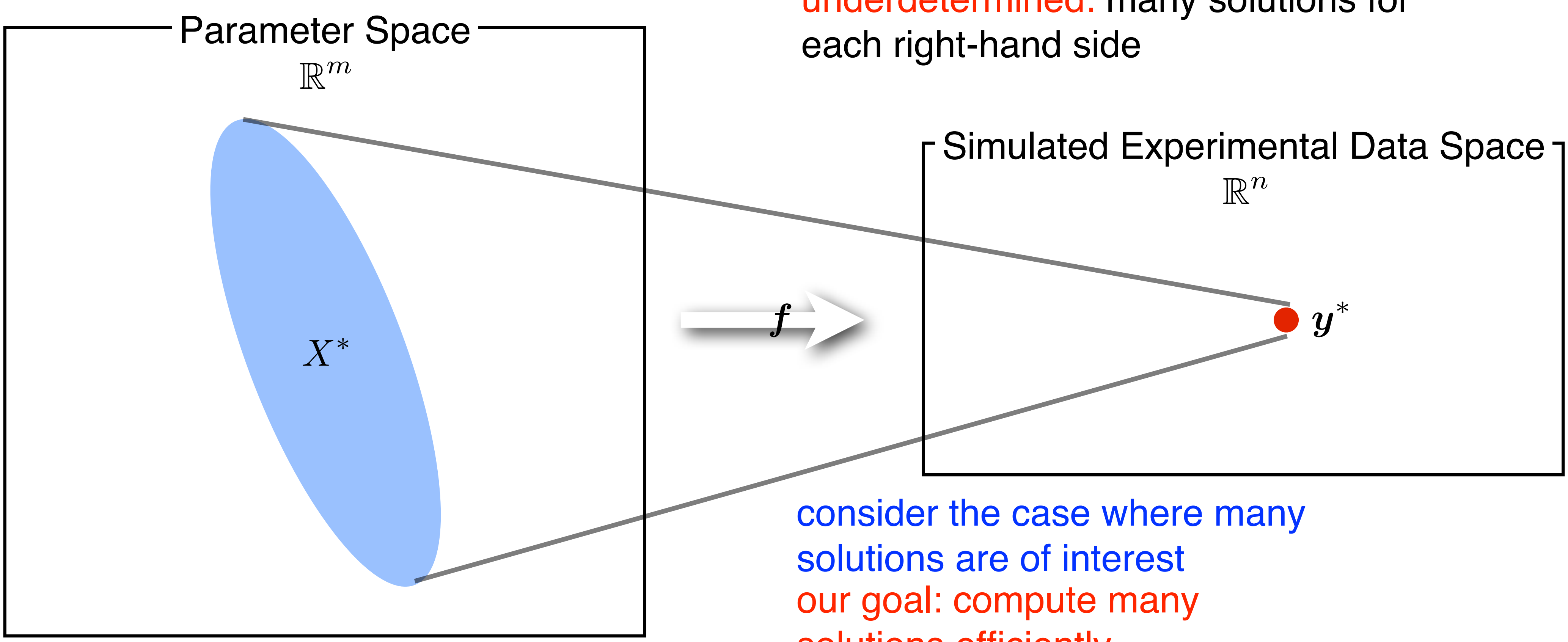
Model Parameter Estimation

underdetermined problem

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}^*$$

$$\boldsymbol{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n \quad m > n$$

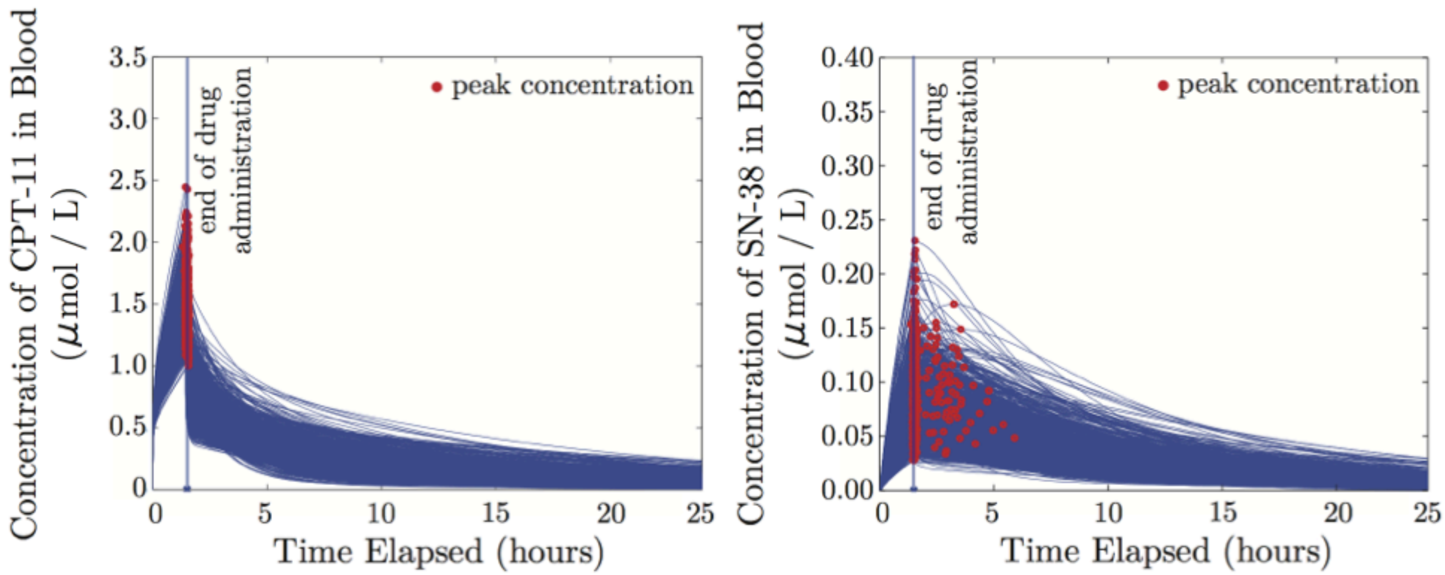
underdetermined: many solutions for each right-hand side



consider the case where many solutions are of interest

our goal: compute many solutions efficiently

Why Multiple Solutions Can Be Useful



(a) Concentration of the anti-cancer drug CPT-11.

(b) Concentration of the metabolite SN-38.

FIG. 1.2. *1,000 model parameter sets found by multiple application of the LM Method.*

Example 1 : Level curve tracing of a rough surface

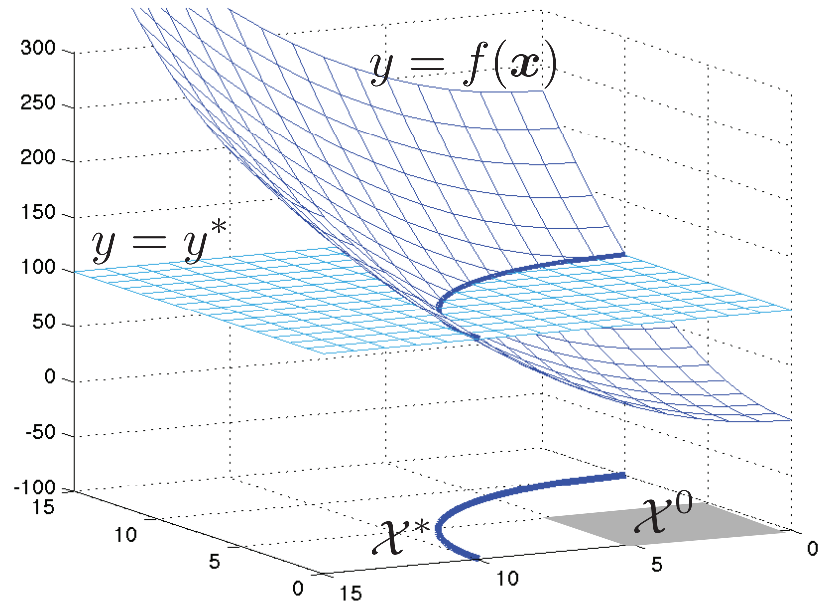
Find a set of 100 points near $\mathcal{X}^0$, s.t.

$$f(\boldsymbol{x}) = y^*$$

where

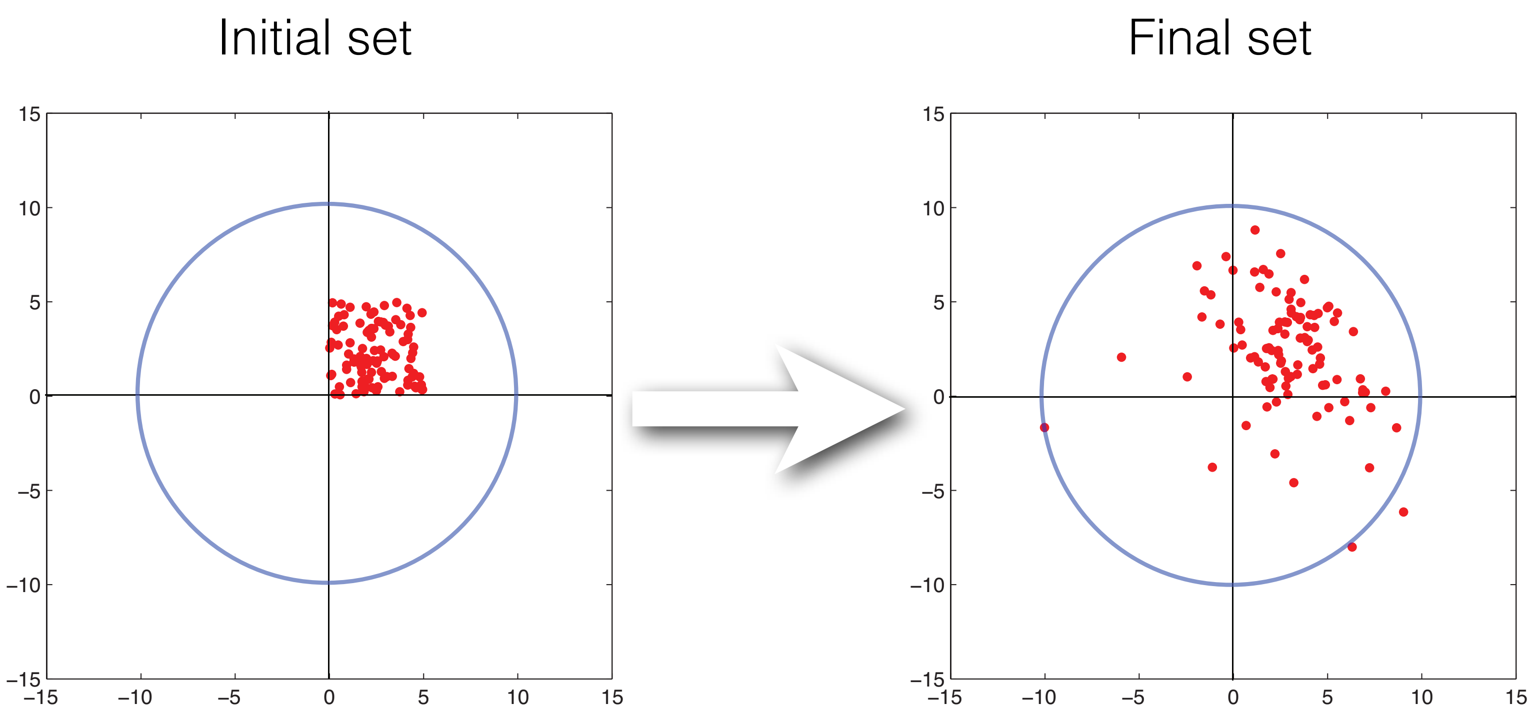
$$f(\boldsymbol{x}) = (x_1^2 + x_2^2) + \sin(10000x_1)\sin(10000x_2)/100$$

$$y^* = 100$$



Levenberg Marquardt method (one-by-one)

For all of the initial points, the algorithm terminated with an error
“Algorithm appears to be converging to a point that is not a root.”

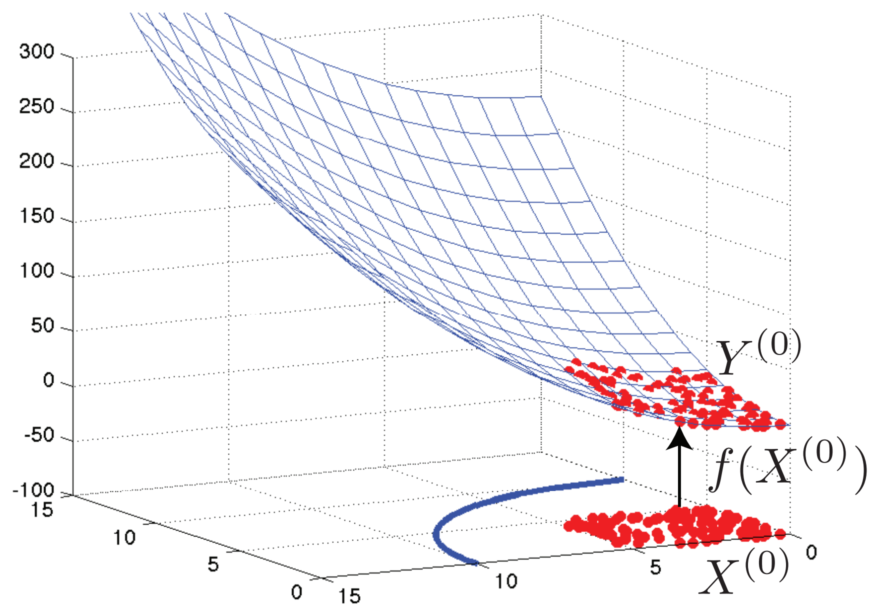


Cluster Newton method for underdetermined problems (collective)

- Stage 1 (Regularized Newton's method applied to a cluster of points)
  - Linear approximation with least squares fitting
    (Least square solution of an overdetermined system of linear equations)
  - Moore-Penrose inverse using the linear approximation
    (Min-norm solution of an underdetermined system of linear equations)

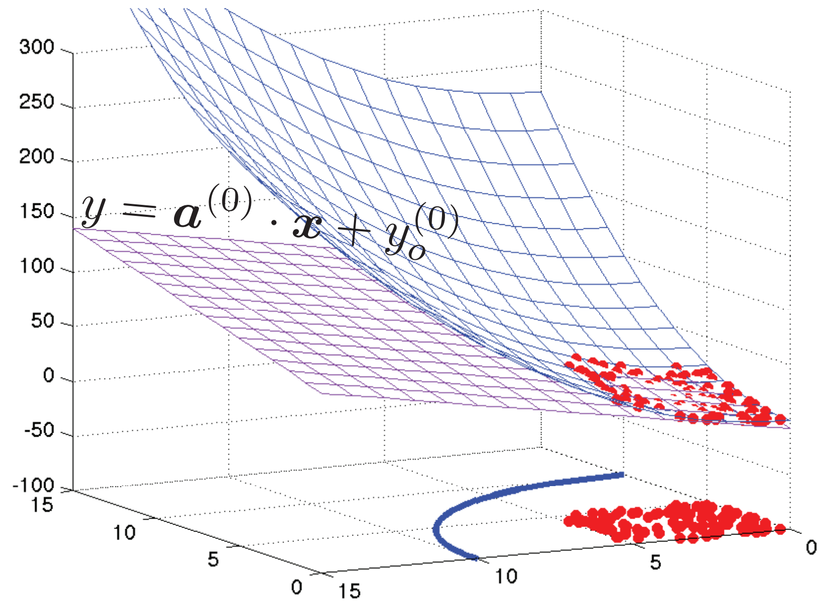
Stage 1

1st iteration step1



Stage 1

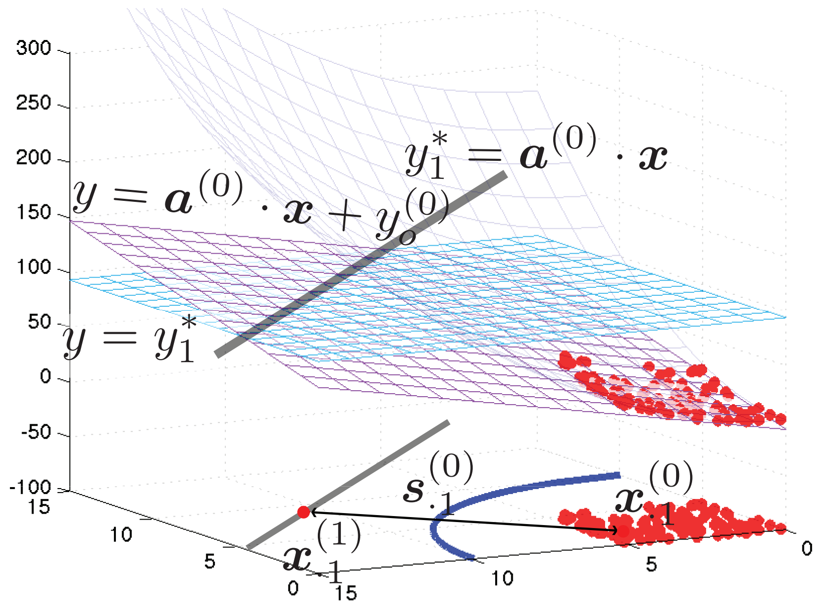
1st iteration step2



compute one collective Jacobian (instead of many Jacobians one-by-one)

Stage 1

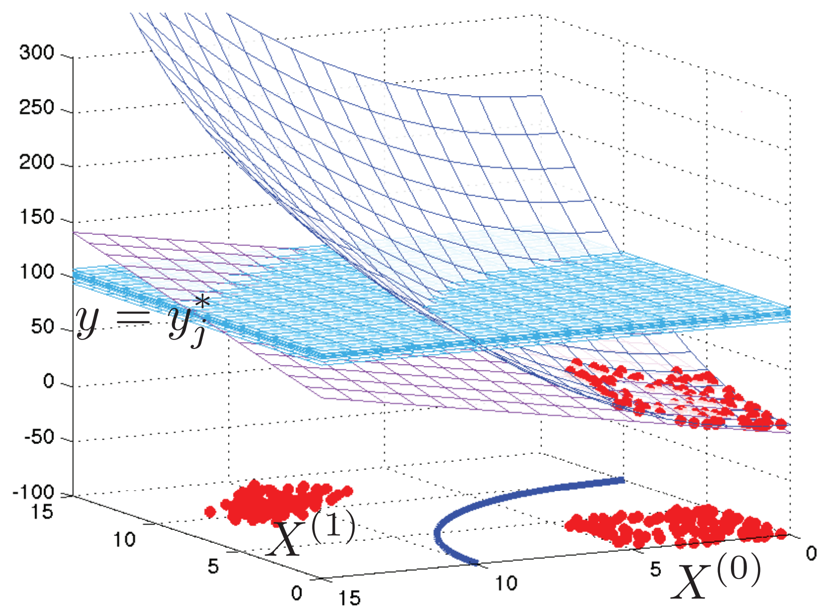
1st iteration step3



find intersection of collective linearization with $y^* = 100$

Stage 1

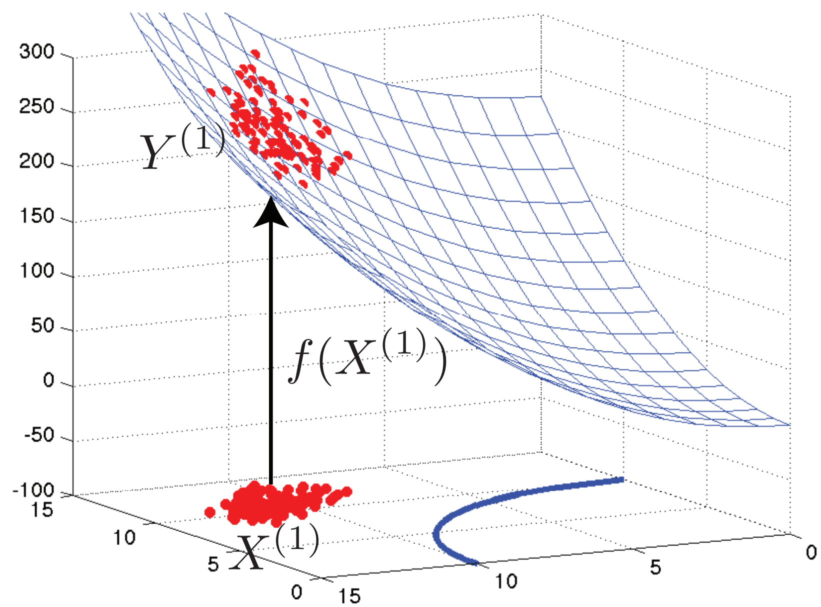
1st iteration step4



for each point, find nearest point on intersection line
(and add a random perturbation to make new points non-collinear)

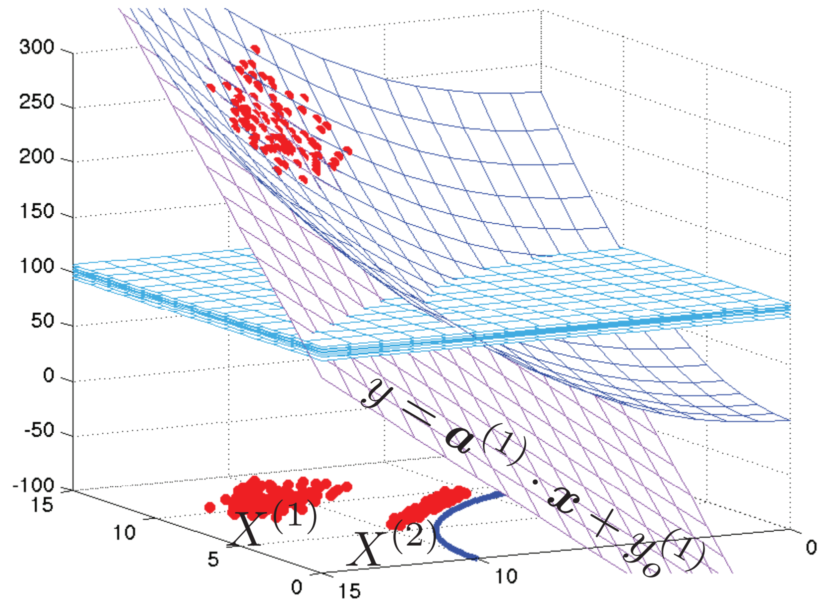
Stage 1

2nd iteration step1

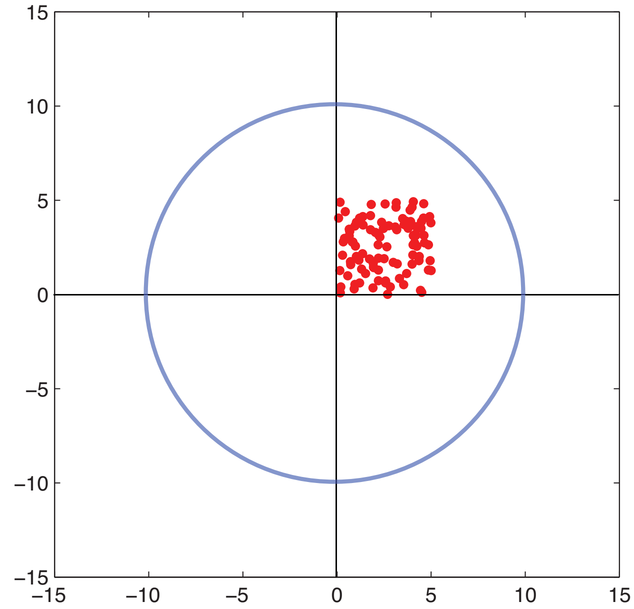


Stage 1

2nd iteration step2-3

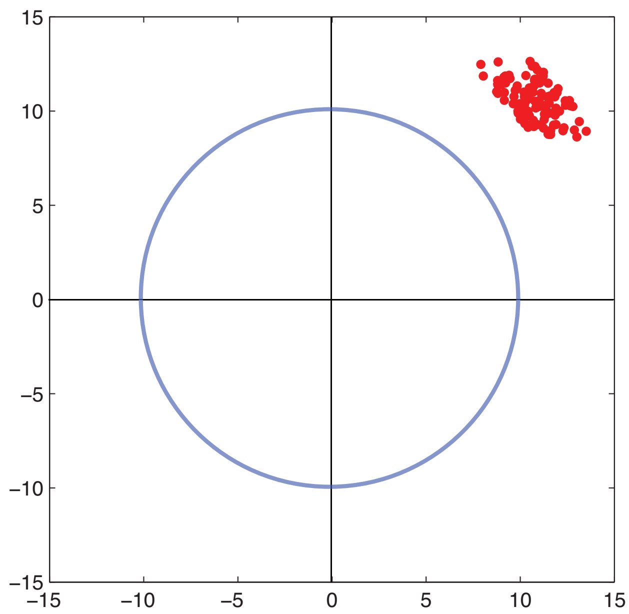


Initial set



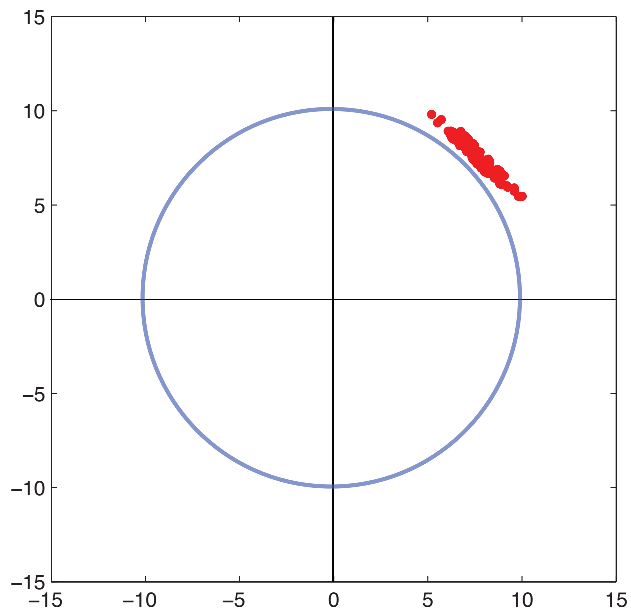
Stage 1

1st iteration



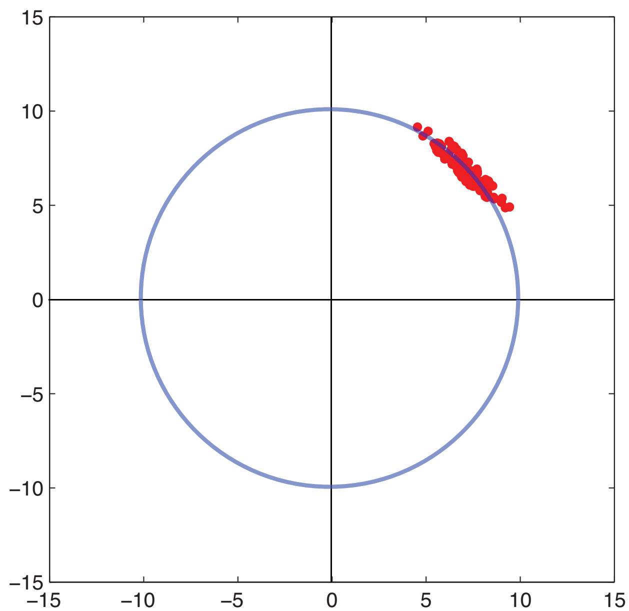
Stage 1

2nd iteration



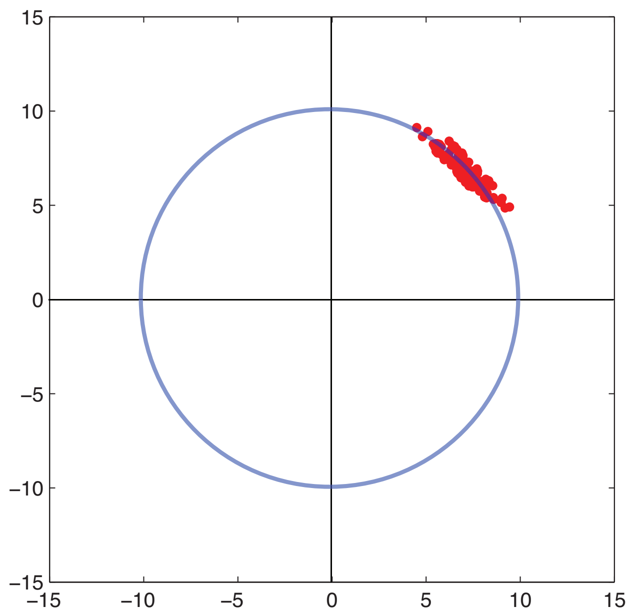
Stage 1

3rd iteration



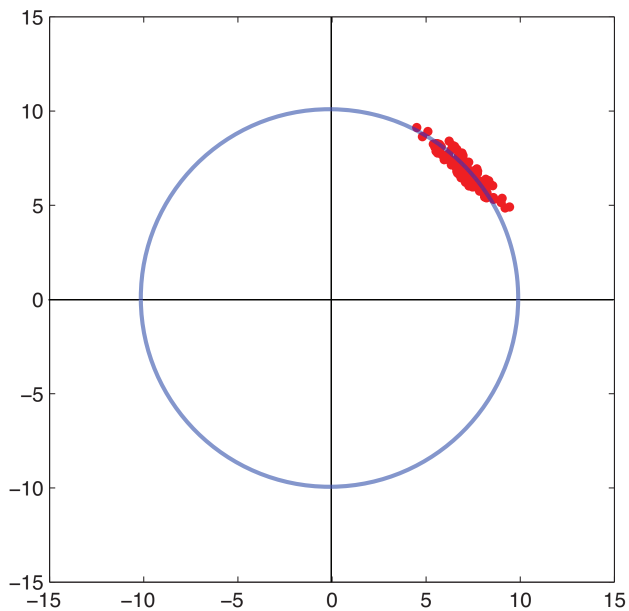
Stage 1

4th iteration



Stage 1

6th iteration

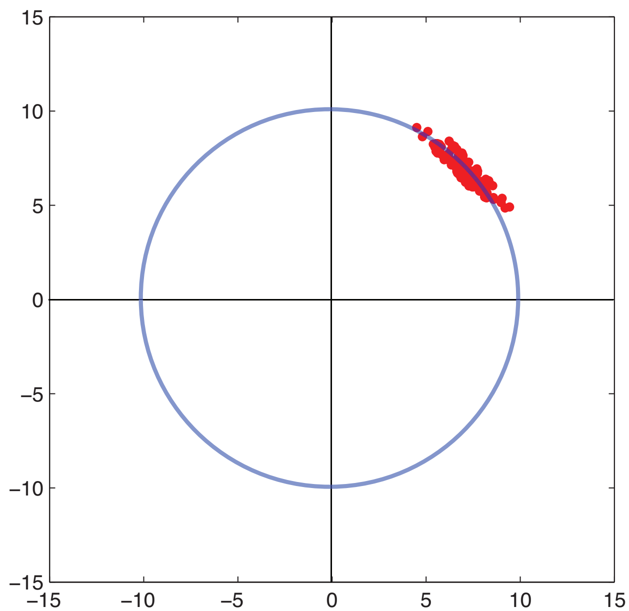


Cluster Newton method for underdetermined problems

- Stage 2 (Broyden’s method, i.e. multi-dimensional secant method)
  - Use collective linear approximation from Stage 1 for initial Jacobian
  - Use the points found by the Stage 1 as initial points

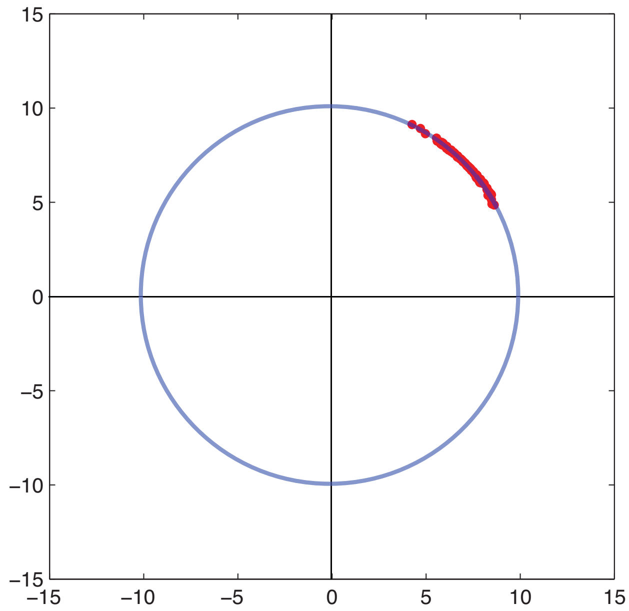
Stage 1

6th iteration



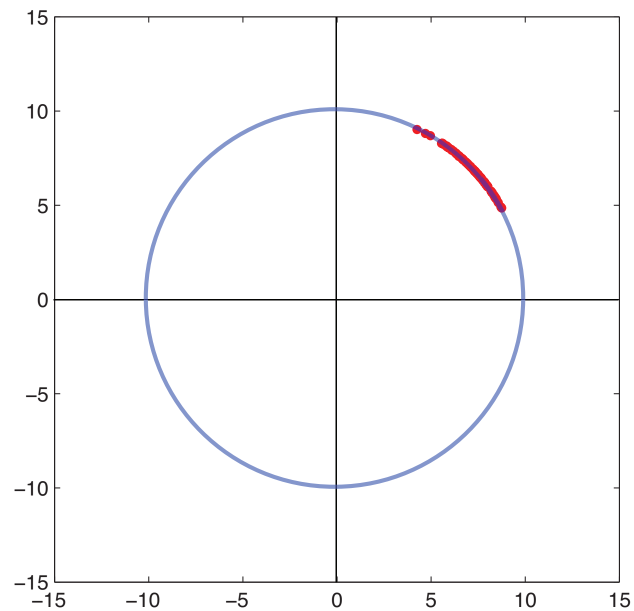
Stage 2

1st iteration



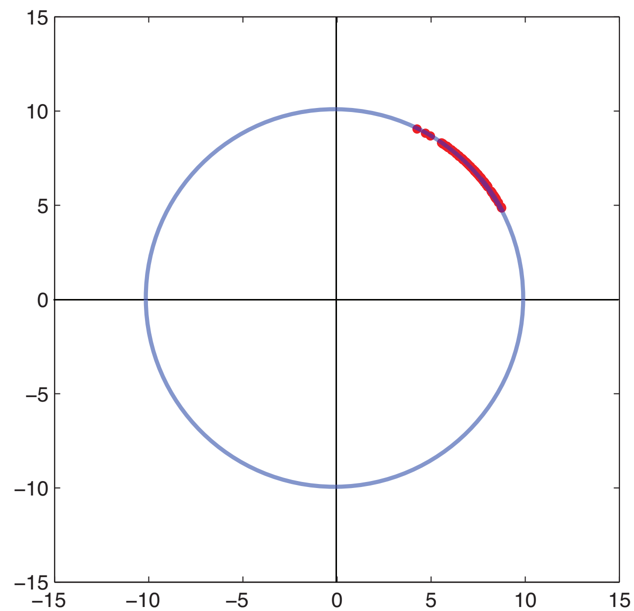
Stage 2

2nd iteration

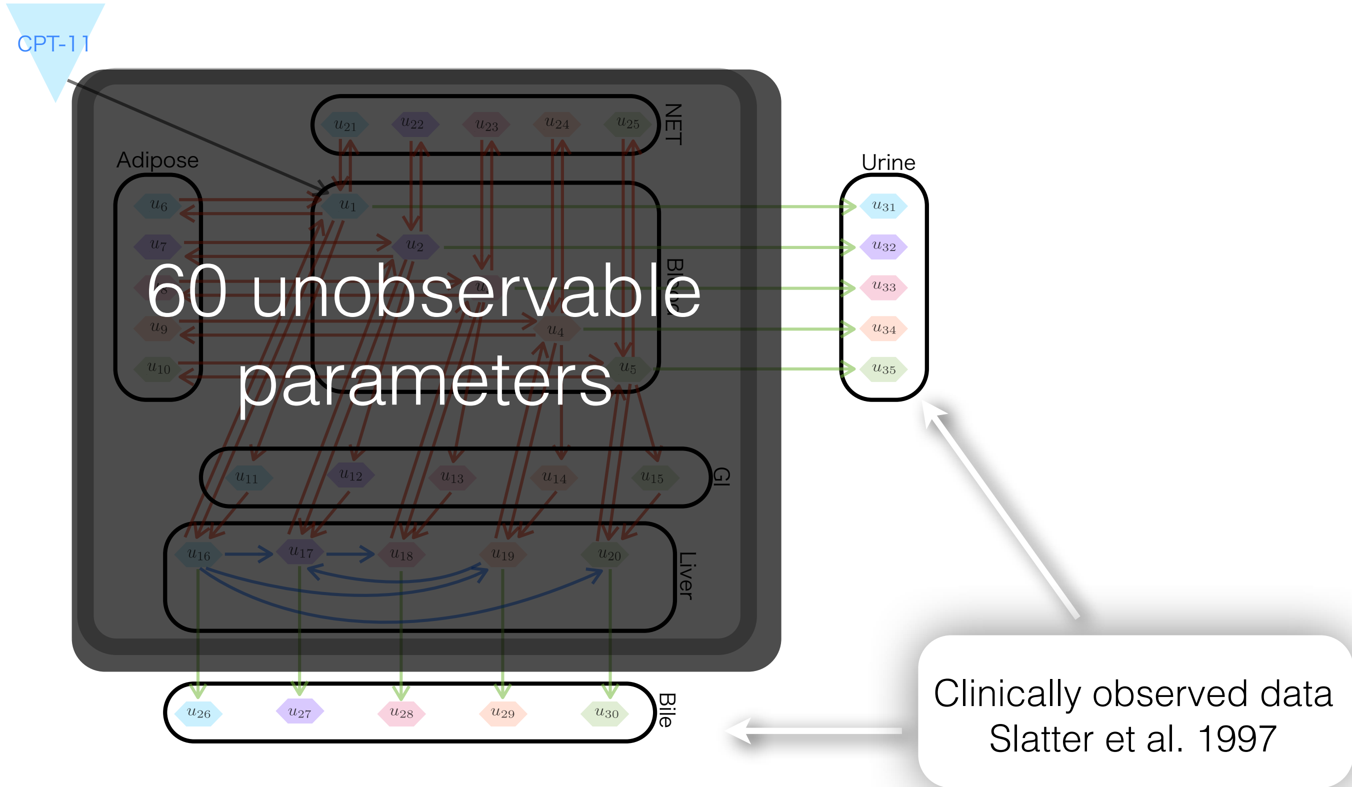


Stage 2

5th iteration



Example 2 : Pharmacokinetics Model (Arikuma et al.)



Estimate model parameters from non-invasive clinical observation,
i.e., find $\boldsymbol{x}$ such that

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}^*$$

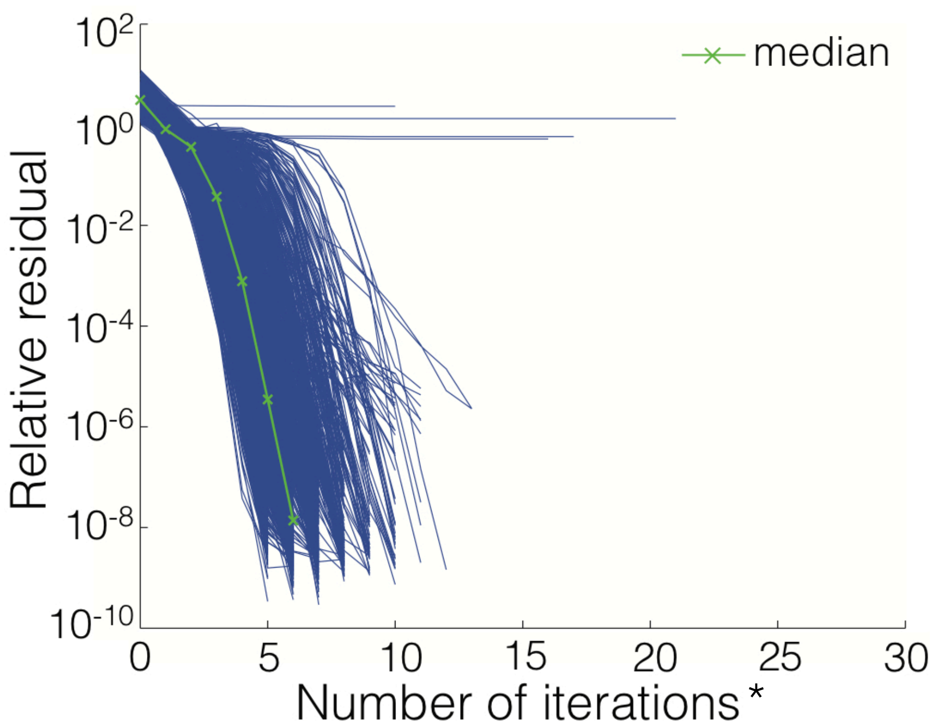
where

$$\boldsymbol{f} : \mathbb{R}^{60} \to \mathbb{R}^{10}$$

(60 parameters, but only 10 observations)

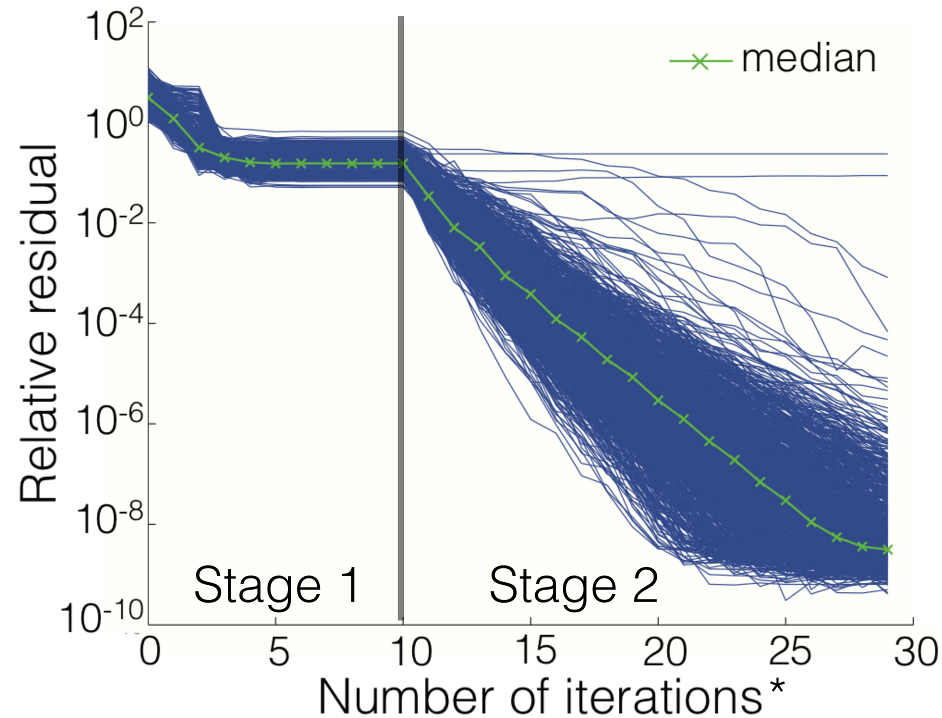
Convergence

Levenberg Marquardt method



* : at least 61 function evaluations / iteration / solution
(469,439 total function evaluations (= ODE solves))

Cluster Newton method

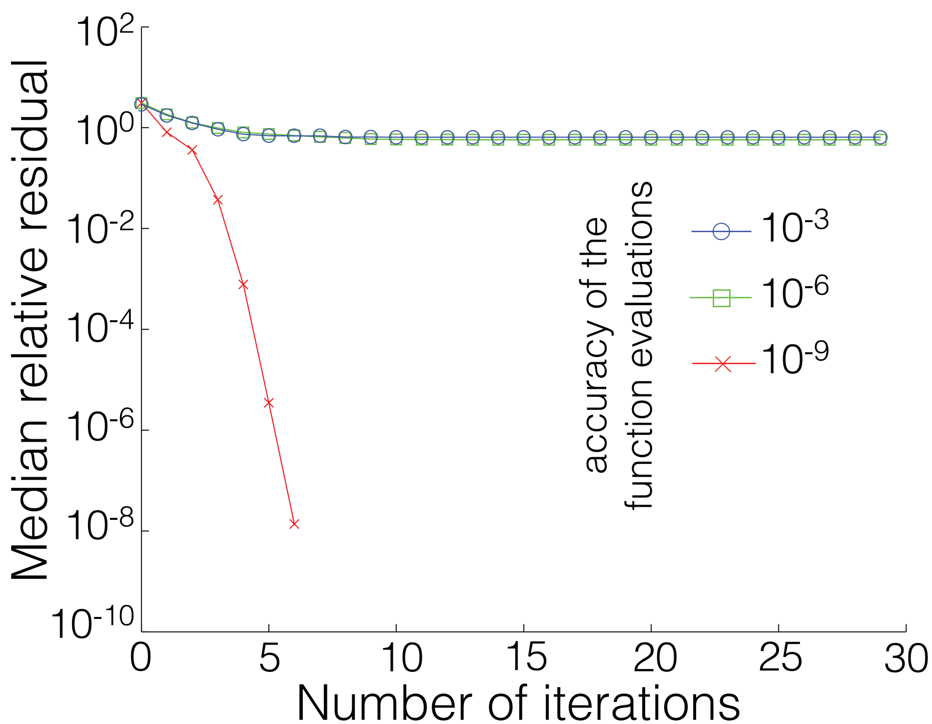


* : only 1 function evaluations / iteration / solution
(30,000 total function evaluations)

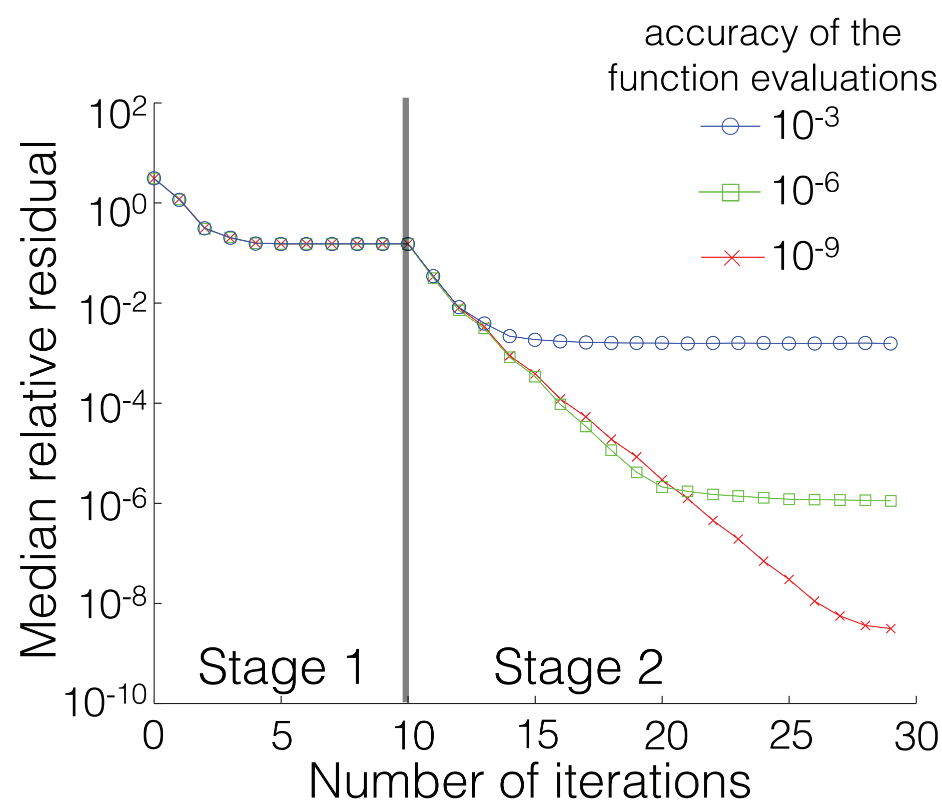
(we seek 1000 solutions)

Sensitivity to the function (ODE solve) evaluation error

Levenberg Marquardt method



Cluster Newton method



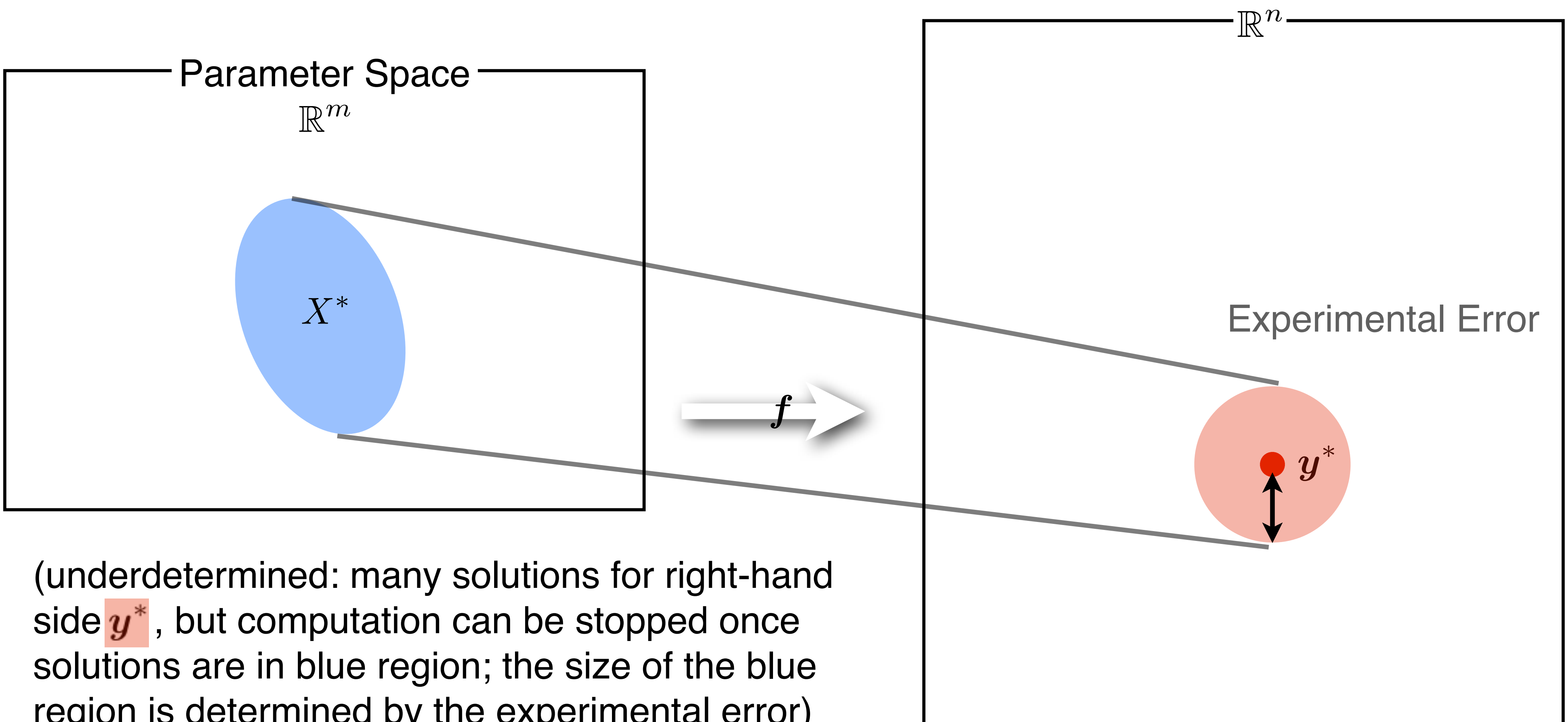
(in practice, target accuracy is determined by experimental error)

Model Parameter Estimation

underdetermined problem

$$\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{y}^*$$

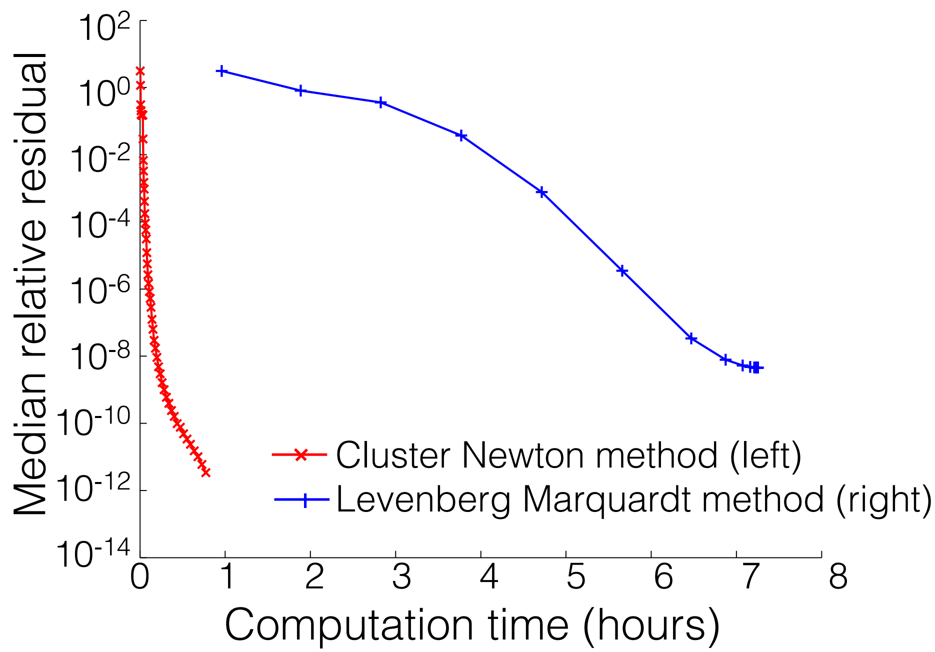
$$\boldsymbol{f} : \mathbb{R}^m \to \mathbb{R}^n \quad m < n$$



(underdetermined: many solutions for right-hand side $\boldsymbol{y}^*$, but computation can be stopped once solutions are in blue region; the size of the blue region is determined by the experimental error)

Computation Speed Comparison :

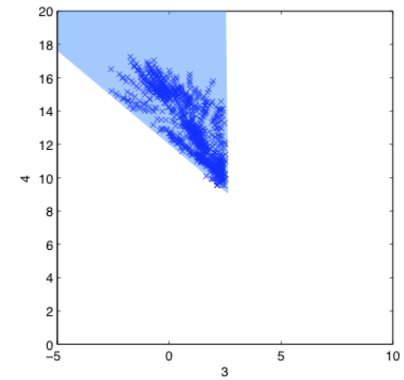
Levenberg-Marquadt method vs. Cluster Newton Method



Overdetermined Parameter Identification Problem

Model Parameter Estimation

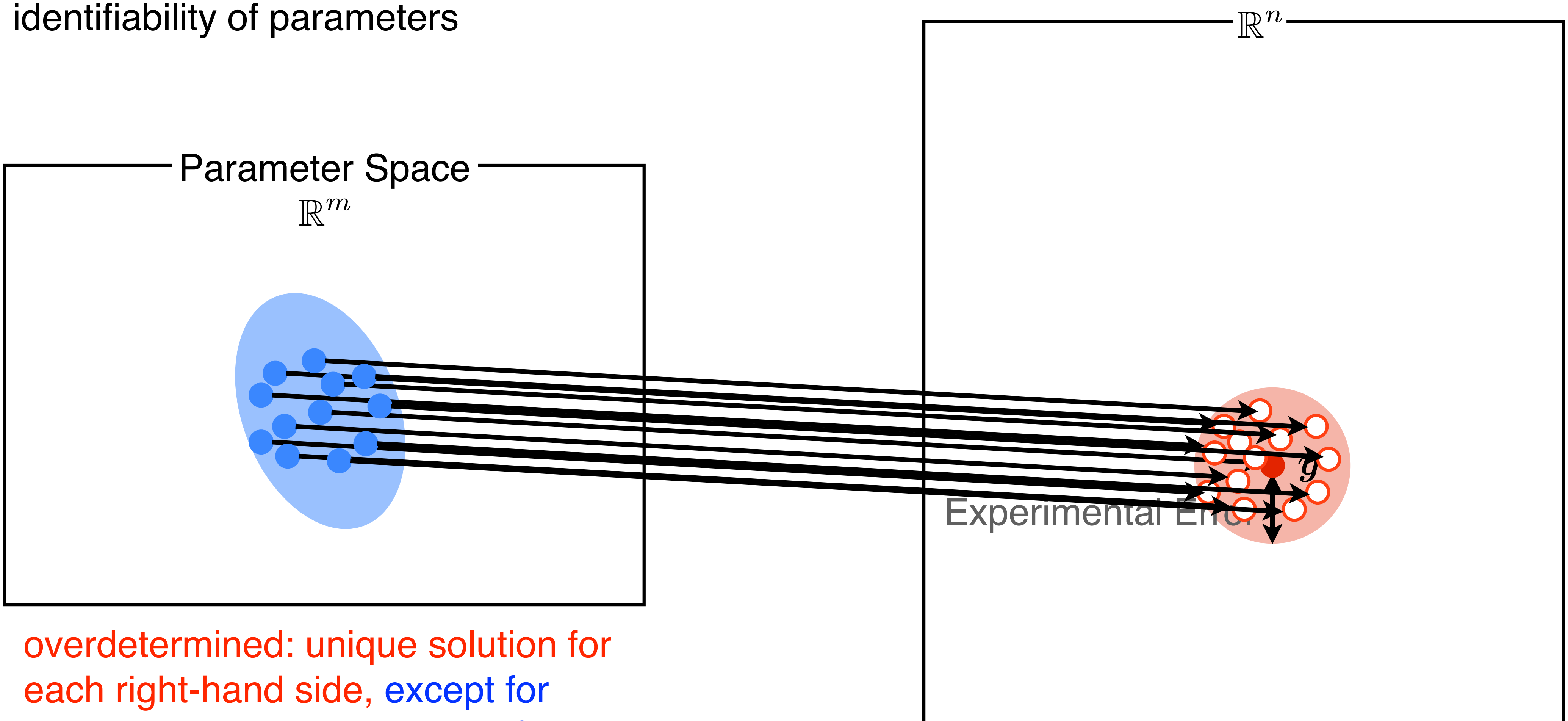
overdetermined problem



$$\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{y}^*$$

overdetermined: finding multiple solutions collectively is often still very useful, to quantify the effect of experimental error, and to investigate identifiability of parameters

$$\boldsymbol{f} : \mathbb{R}^m \to \mathbb{R}^n \qquad m < n$$



overdetermined: unique solution for each right-hand side, except for parameters that are not identifiable

Example 3 : Influenza Kinetics model (Baccam et al.)

Finitely Parameterized Mathematical Model

$$\frac{du_1}{dt} = -x_1 u_1 u_4$$

$$\frac{du_2}{dt} = x_1 u_1 u_4 - \frac{1}{x_2} u_2$$

$$\frac{du_3}{dt} = \frac{1}{x_2} u_2 - \frac{1}{x_3} u_3$$

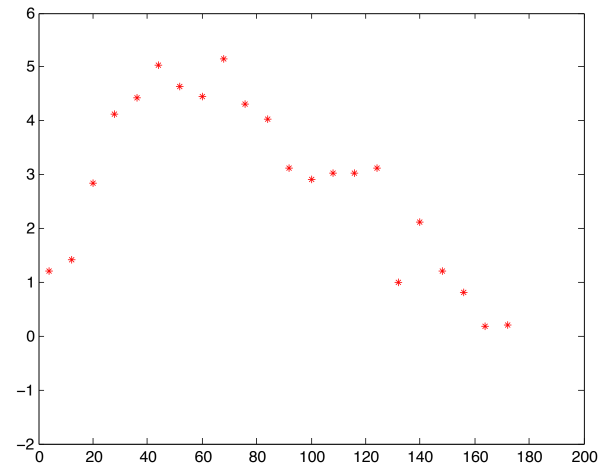
$$\frac{du_4}{dt} = \frac{x_4}{x_5} u_3 - x_6 u_4$$

$$u_1(t=0) = x_5$$
$$u_2(t=0) = 0$$
$$u_3(t=0) = 0$$
$$u_4(t=0) = x_7$$

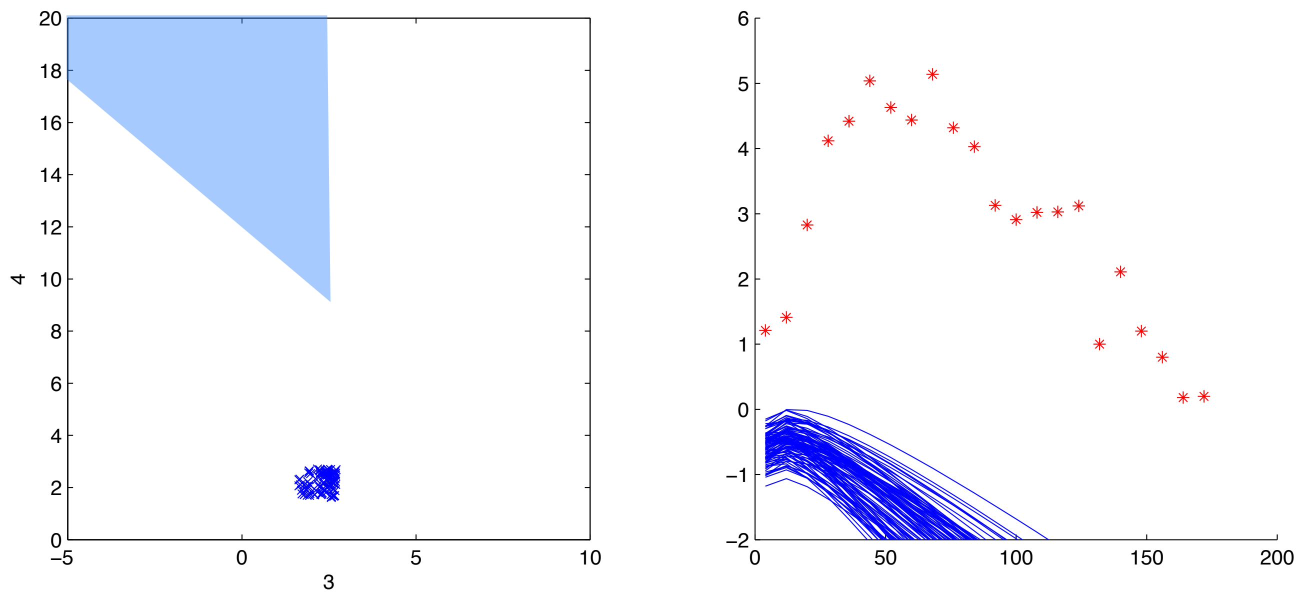
$$\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{y}^*$$

$$\boldsymbol{f} : \mathbb{R}^7 \to \mathbb{R}^{22}$$

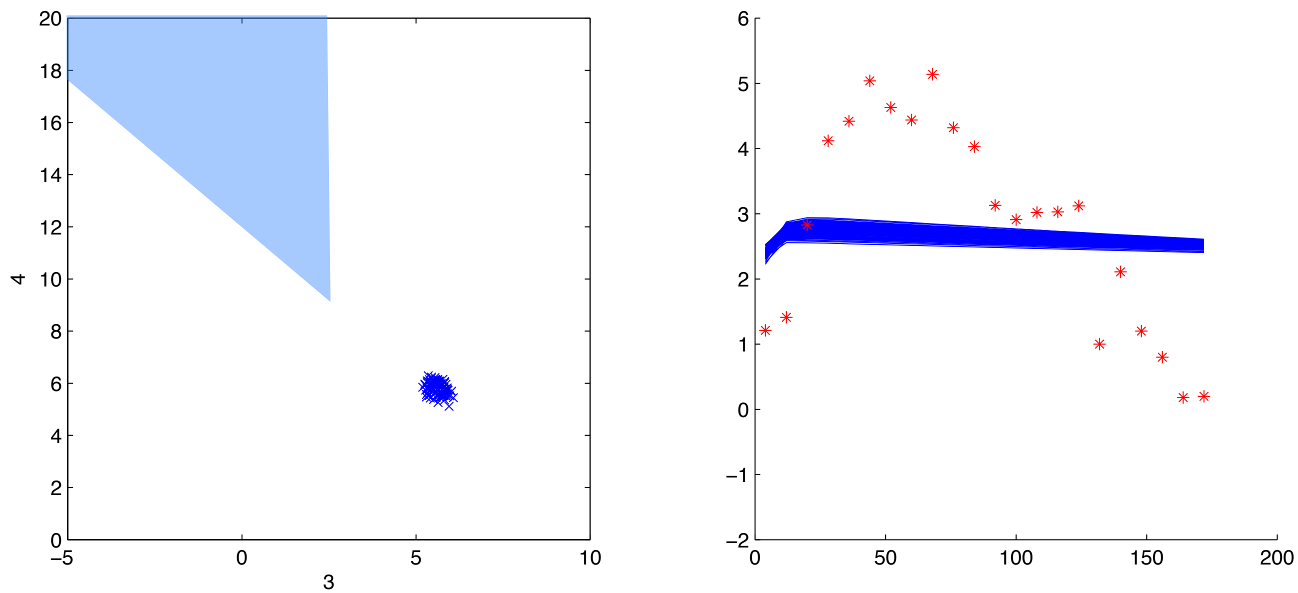
Experimental Data



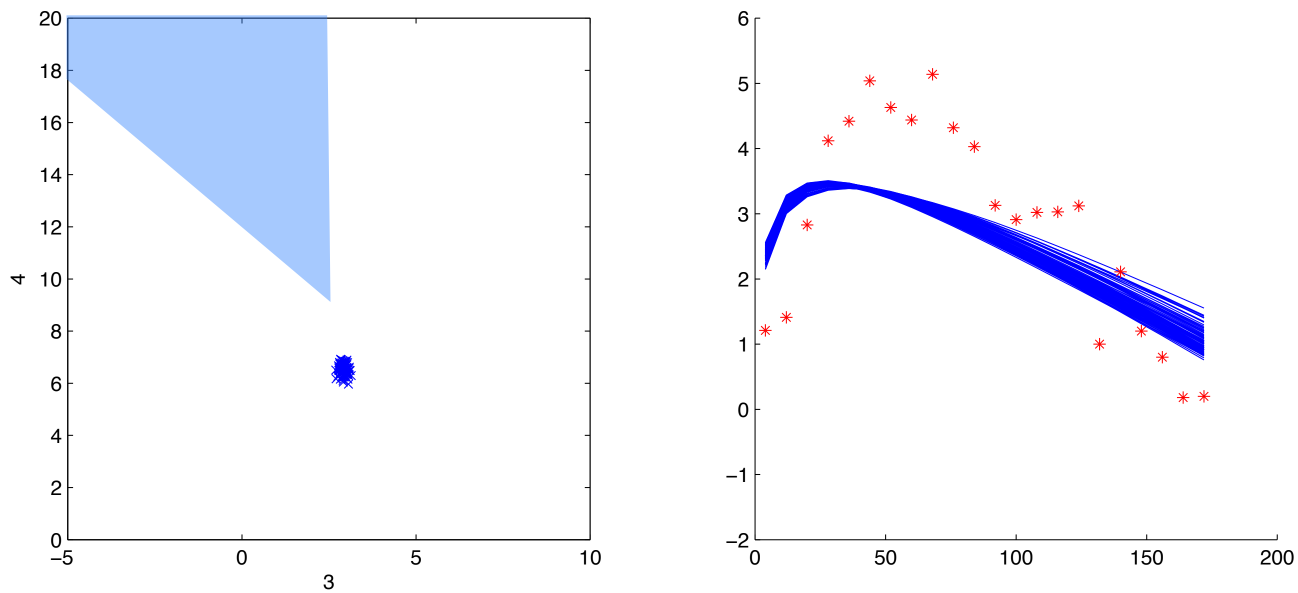
overdetermined, but some parameters are not identifiable

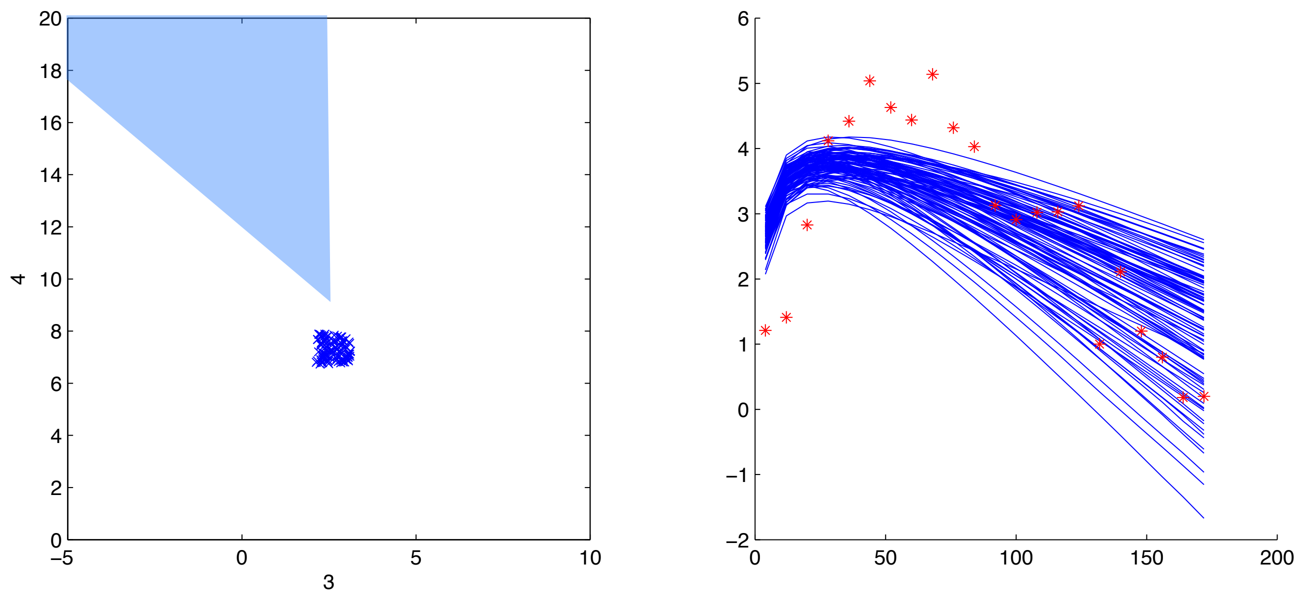


modified Cluster-Newton method can be used to find multiple solutions ‘covering’ the blue region efficiently

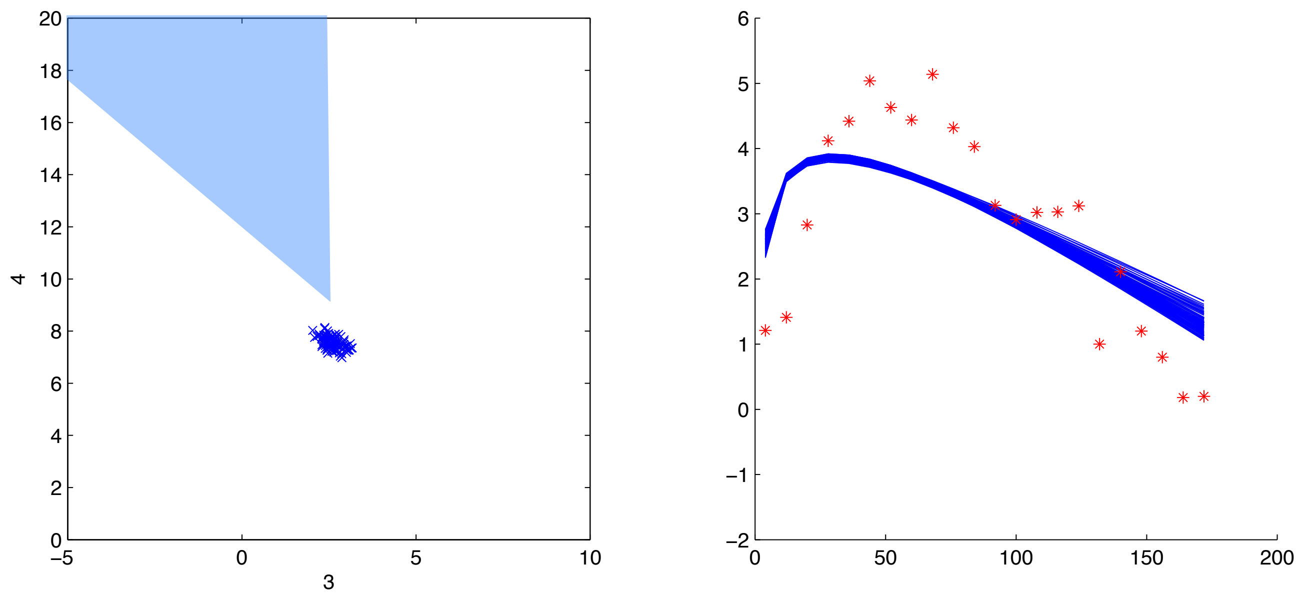


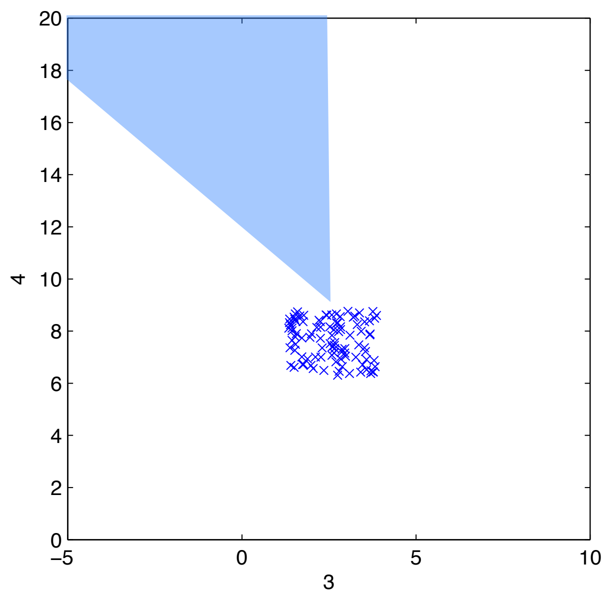
20
18
16
14
12
10
8
6
4
2
0
−5
0
5
10
3
4
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200


20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

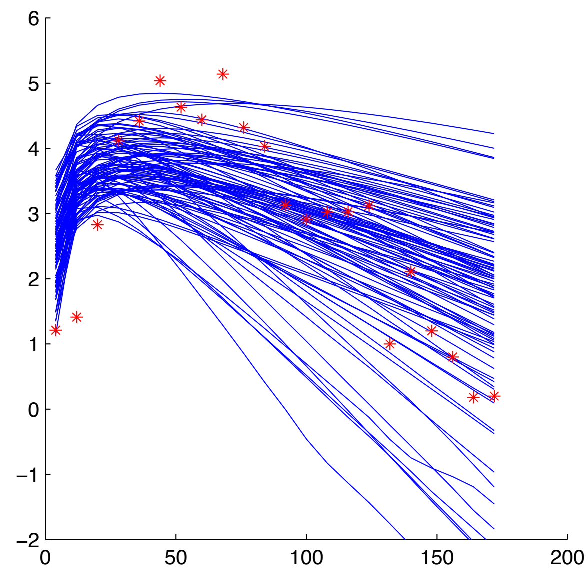


20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

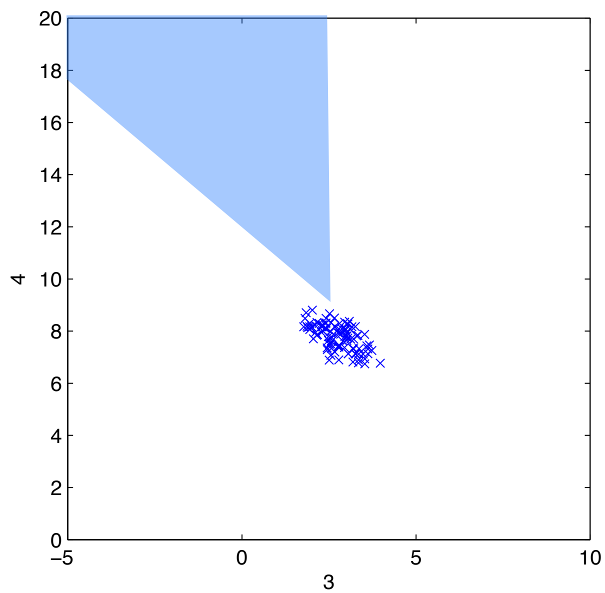

20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200



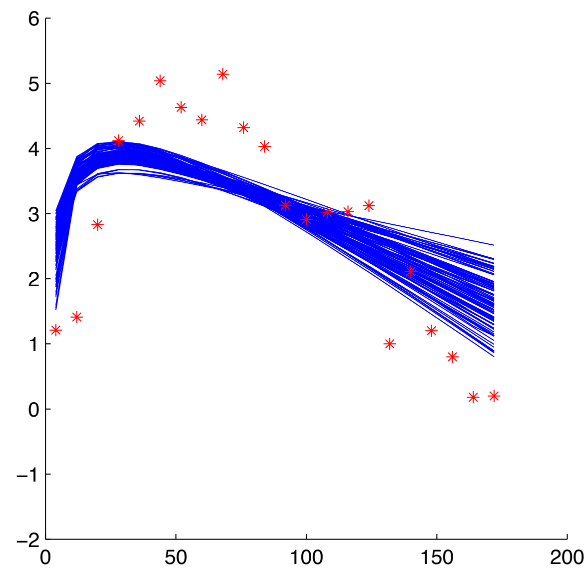
20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3



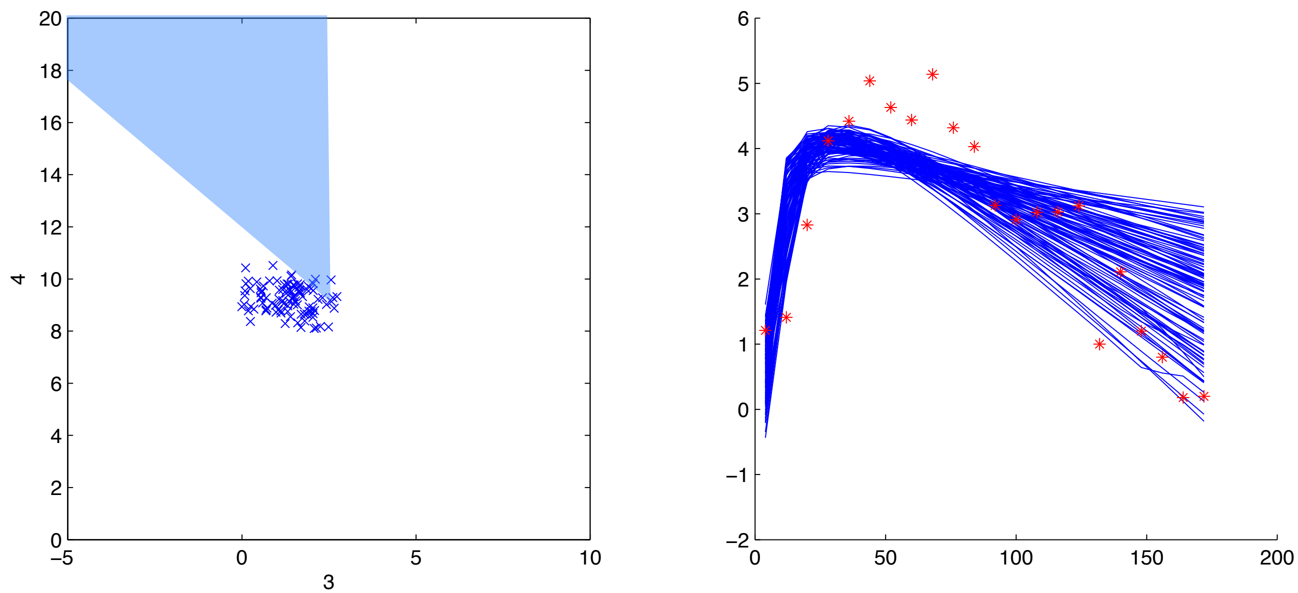
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200



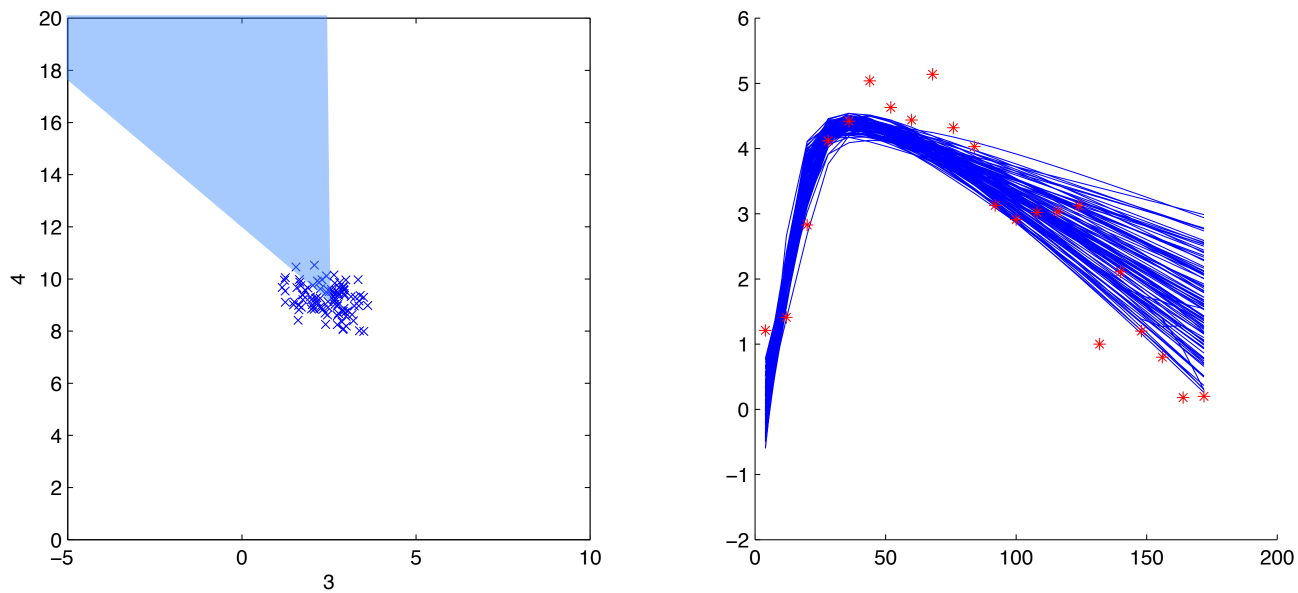
20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3



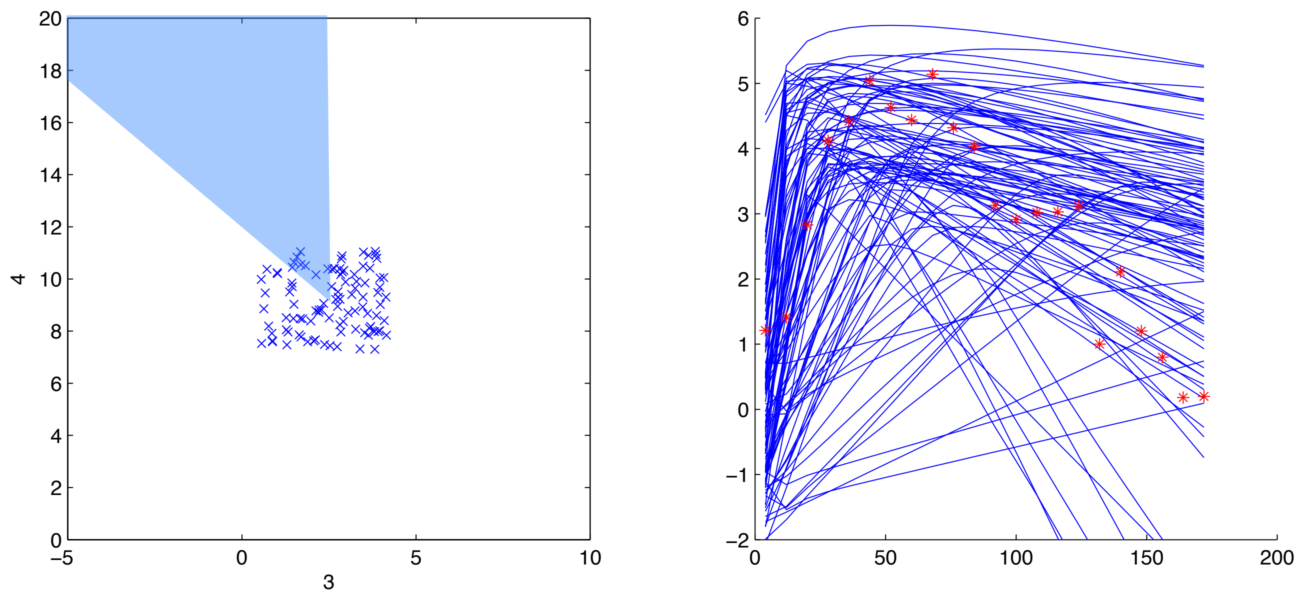
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200



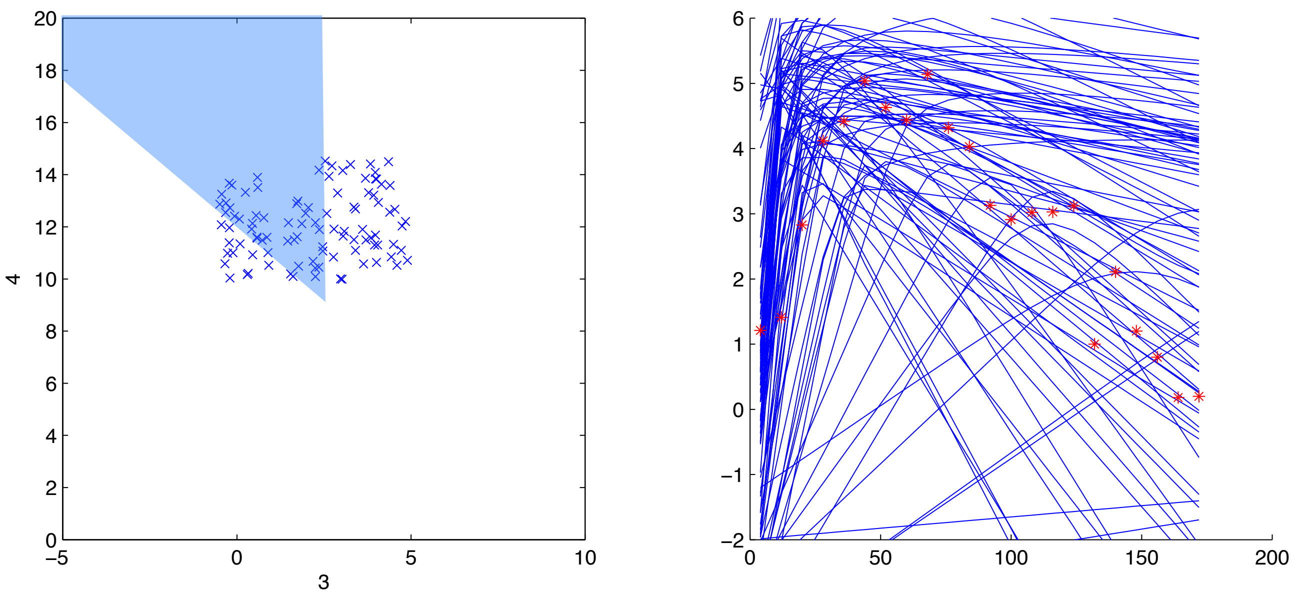
20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

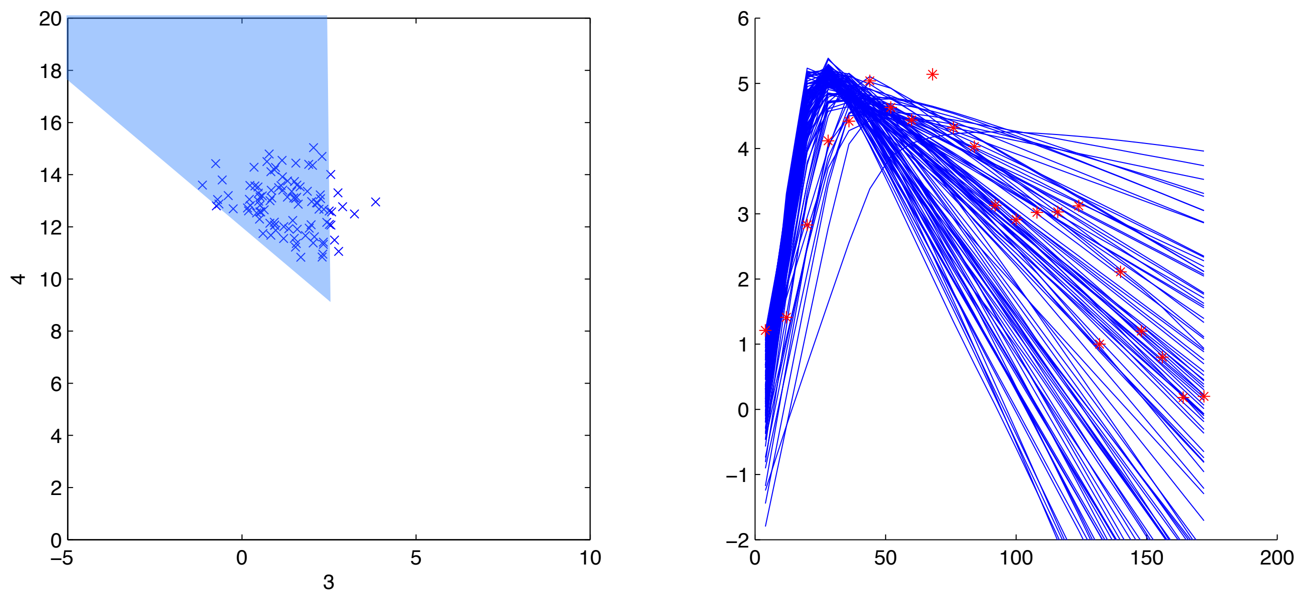


20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

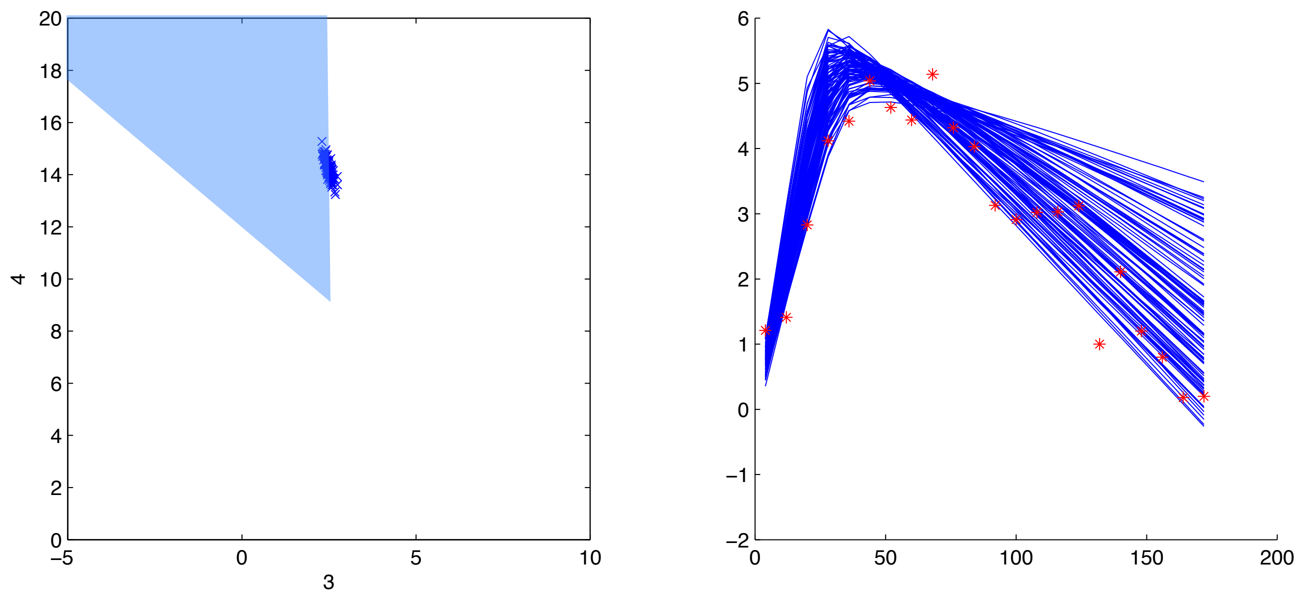


20
18
16
14
12
10
8
6
4
2
0
−5
0
5
10
3
4
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

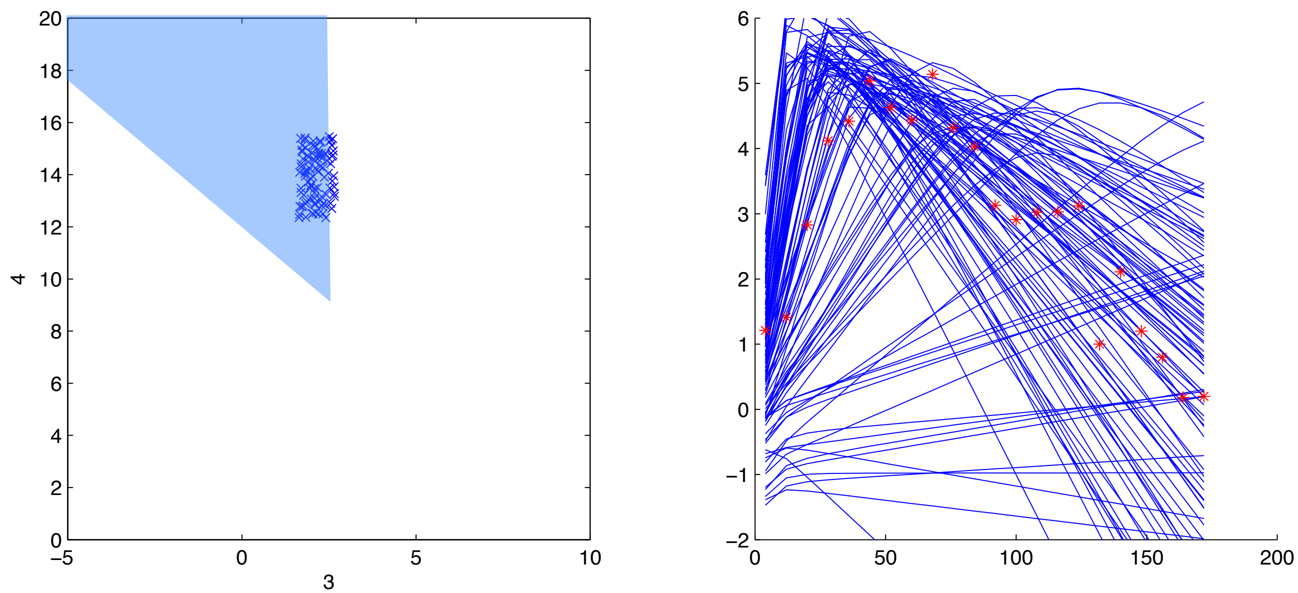

20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

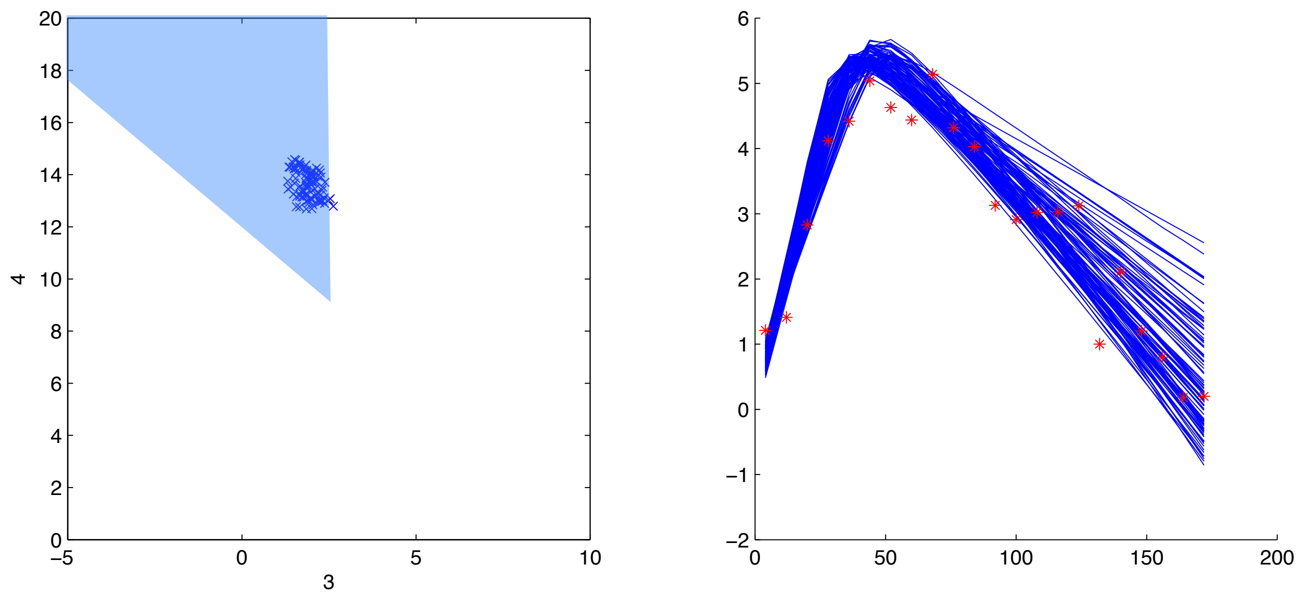


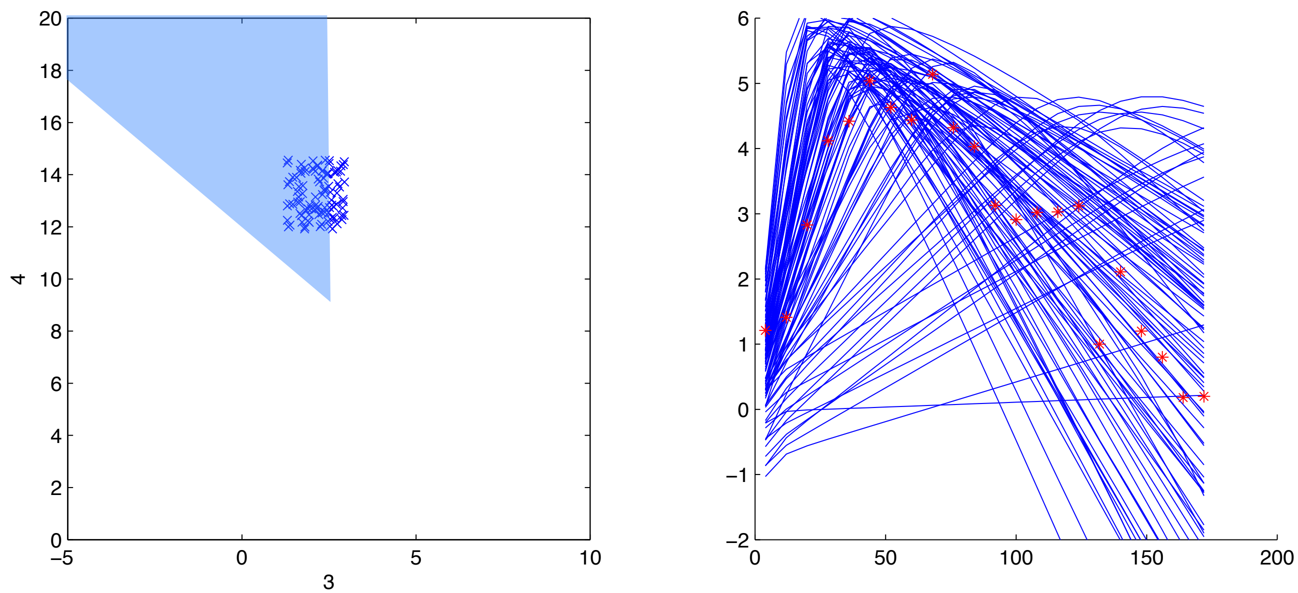
20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200



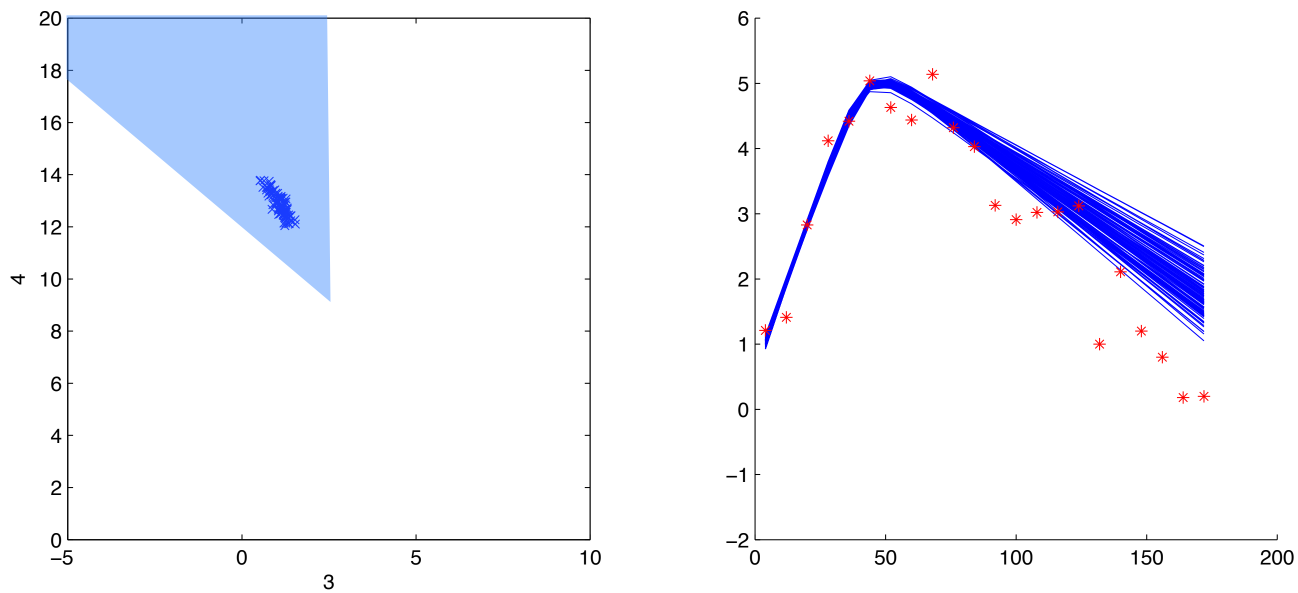
20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200


20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

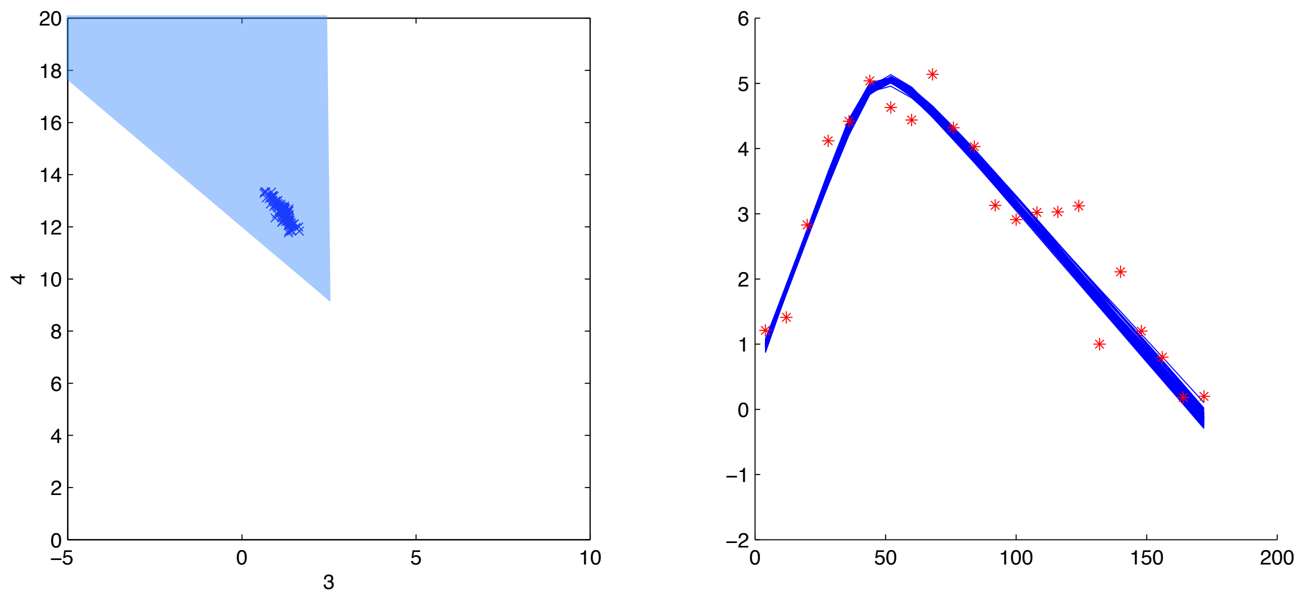

20
18
16
14
12
10
8
6
4
2
0
−5
0
5
10
3
4
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

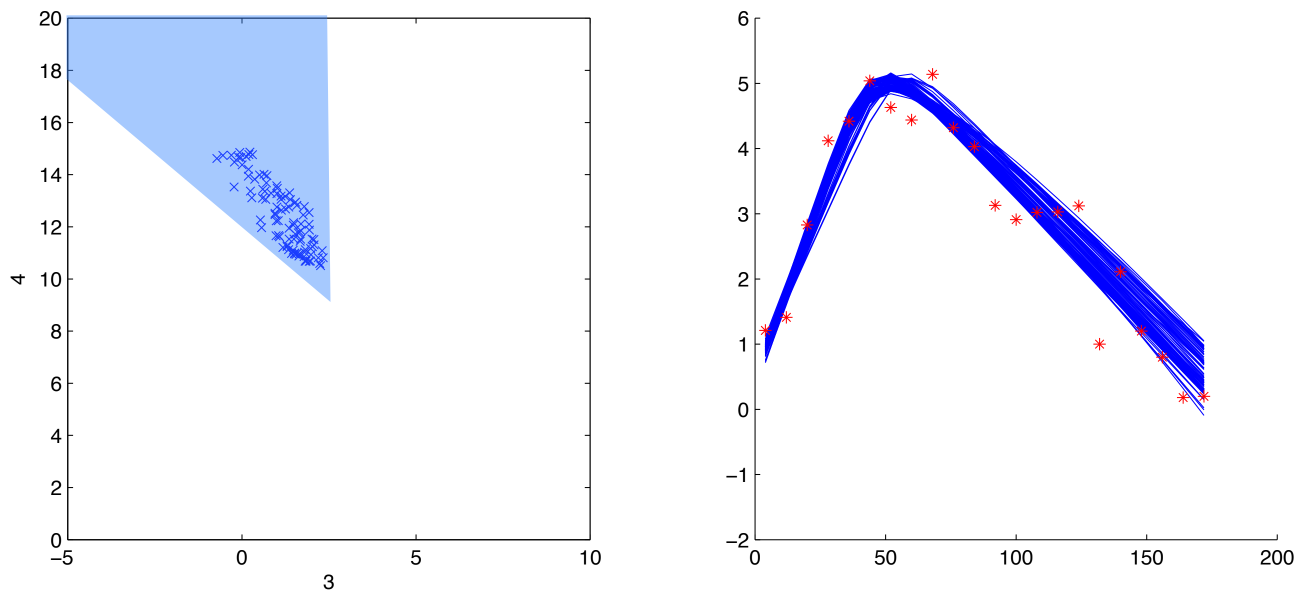


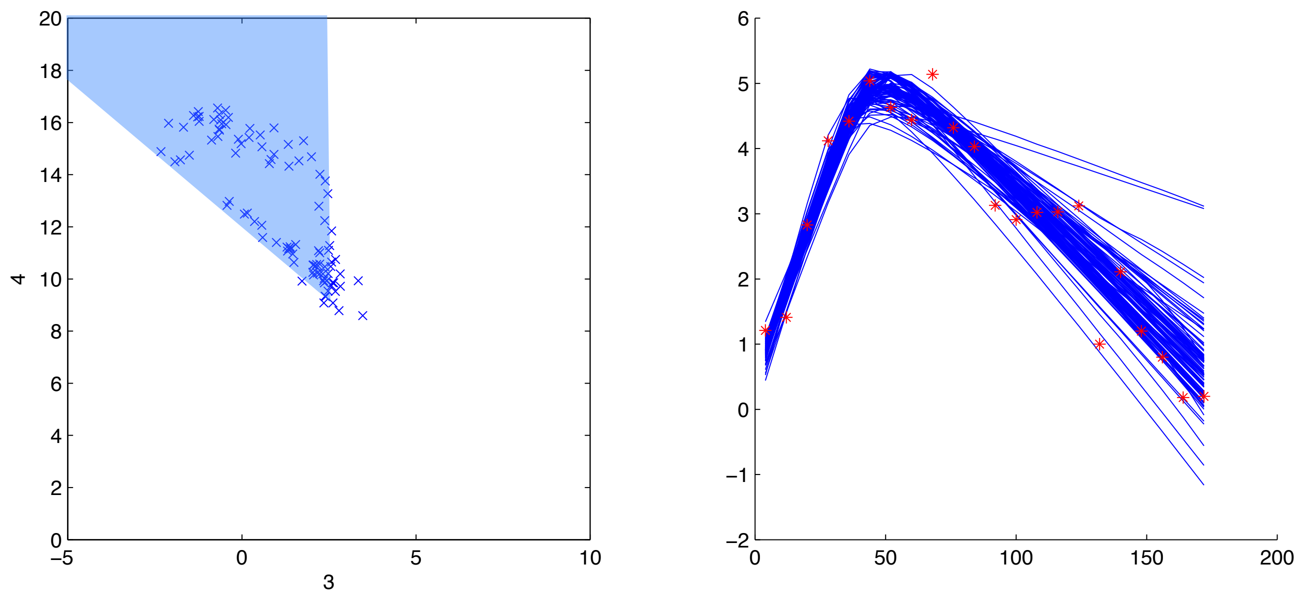
20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200



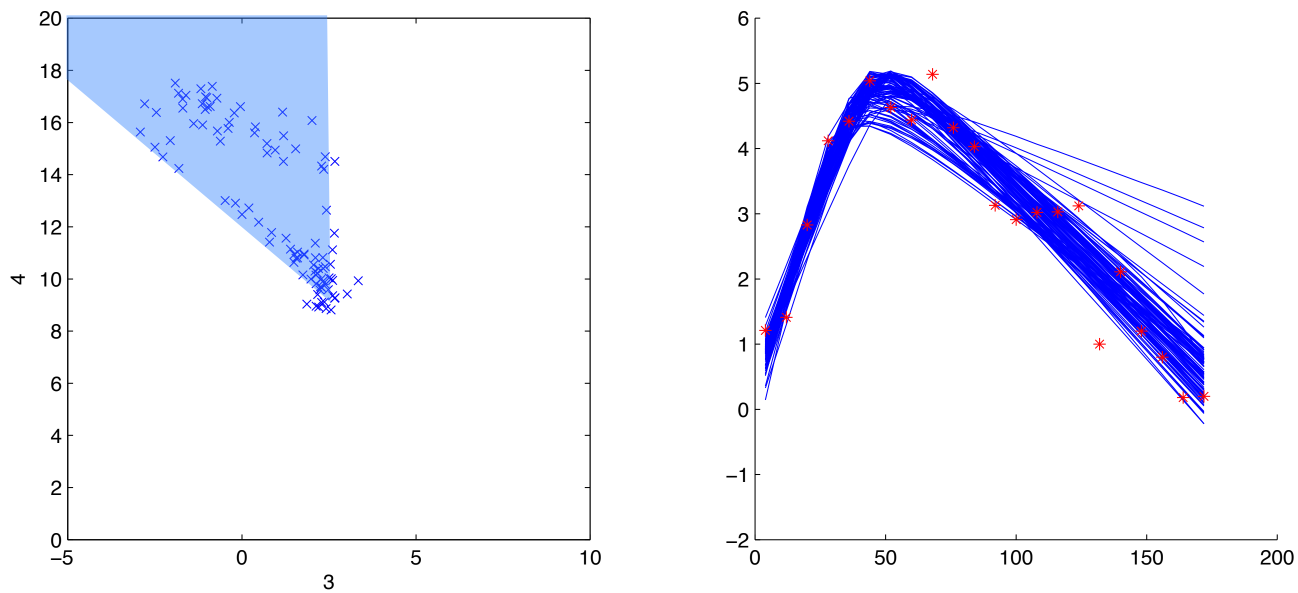
20
18
16
14
12
10
8
6
4
2
0
−5
0
5
10
3
4
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200


20
18
16
14
12
10
8
6
4
2
0
4
−5
0
5
10
3
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200


20
18
16
14
12
10
8
6
4
2
0
−5
0
5
10
3
4
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200

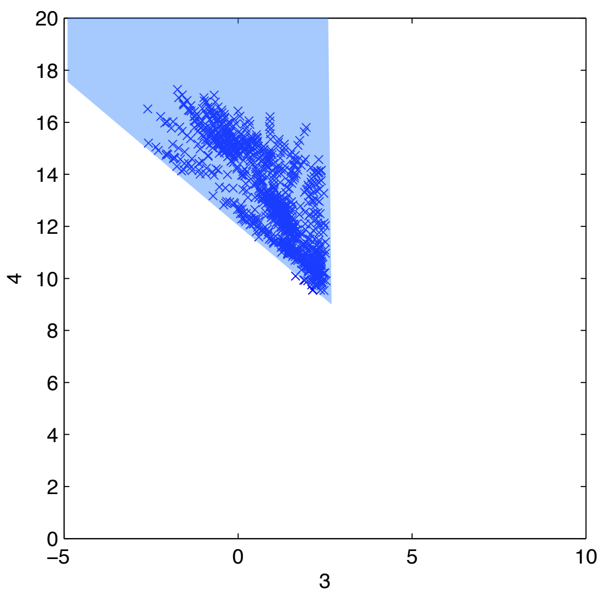


20
18
16
14
12
10
8
6
4
2
0
−5
0
5
10
3
4
6
5
4
3
2
1
0
−1
−2
0
50
100
150
200



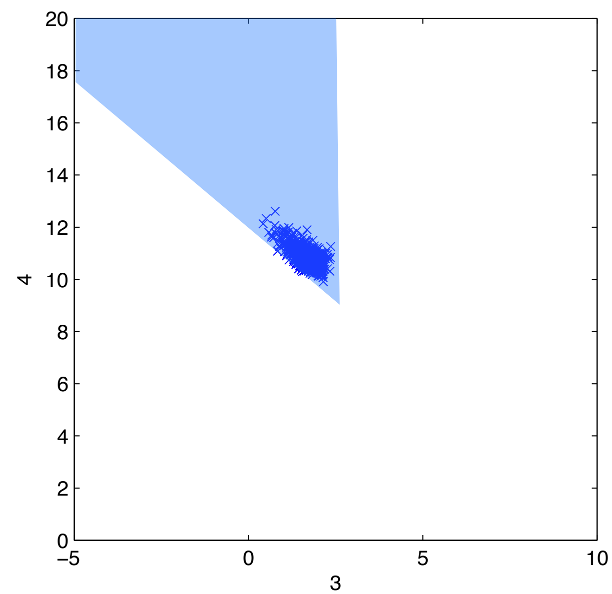
3
4

Cluster Newton method



After **7,885** function (= ODE) evaluations

Genetic Algorithm + Monte Carlo method



After **9,108 + 56,957** function evaluations

Conclusion

Cluster Newton method is a computationally efficient algorithm
to find multiple reasonable parameters of finitely parameterized mathematical models.

It is efficient and robust due to collective fitting of the Jacobian.

[1] Yasunori Aoki, Ken Hayami, Hans De Sterck and Akihiko Konagaya, Cluster Newton Method for Sampling Multiple Solutions of an Underdetermined Inverse Problem: Parameter Identification for Pharmacokinetics, Technical Report NII-2011-002E, National Institute of Informatics ,Tokyo, Japan (2011).

[2] Yasunori Aoki, Ken Hayami, Hans De Sterck, and Akihiko Konagaya, 'Cluster Newton Method for Sampling Multiple Solutions of Underdetermined Inverse Problems: Application to a Parameter Identification Problem in Pharmacokinetics', SIAM Journal on Scientific Computing, accepted, 2013.