Estimation of Parameters in Models for Dynamics Processes - Formulations, Numerical Methods, Applications

Johannes P. Schlöder

Interdisciplinary Center for Scientific Computing (IWR)
Heidelberg University

based on joint work with

A. Altmann-Dieses, H. G. Bock, K. Hatz, S. Lenz, and E. Kostina

Systems Control and Optimization Laboratory
IMTEK, Faculty of Engineering, University of Freiburg, November 26, 2018

Overview

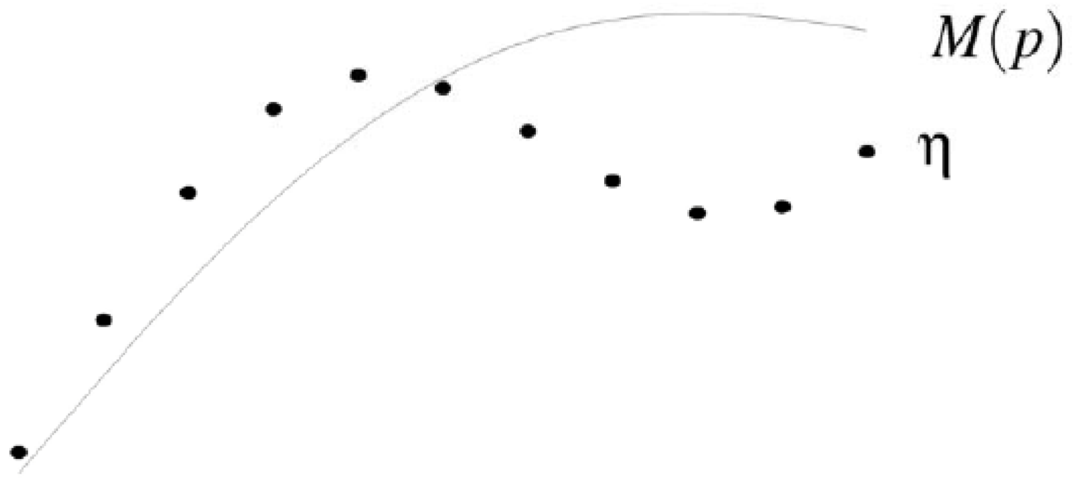
- Parameter Estimation Problems
  - Example: Urethane Reaction
  - Differential Equation and Optimal Control Models & Data
  - Optimization Boundary Value Problems

- Structure Exploiting Numerical Methods for
  Optimization Boundary Value Problems

  - The Direct Multiple Shooting Method for Parameter Estimation
  - The Generalized Gauss Newton Method
  - Assessment of the Statistical Error of the Parameter Estimates

- "Proof of Concept" and "Real World“ Applications

Parameter Estimation: Match Model to Data



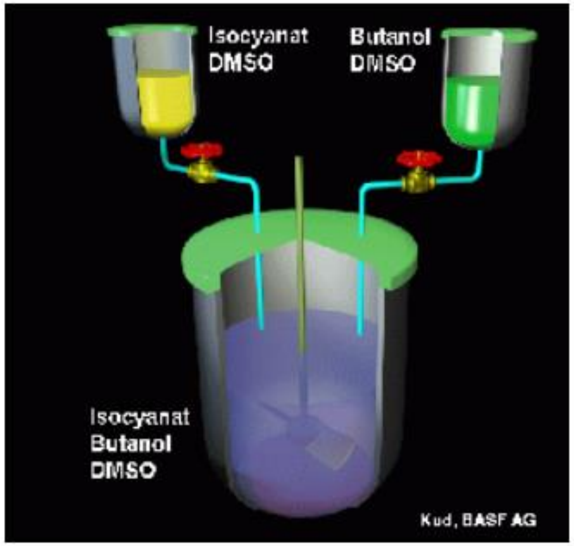
M(p) **model response** —
η **data** •

Classroom Example: The Reaction of Urethane

The Reaction of Urethane

$$
\begin{aligned}
A + B &\longrightarrow C \\
A + C &\rightleftharpoons D \\
3\,A &\longrightarrow E
\end{aligned}
$$

A: isocyanate B: butanol
C: urethane D: allophanate
E: isocyanurate L: solvent DMSO



$$\vdots$$

$$\dot{n}_E = V \cdot k_{ref_4} \cdot \exp\left(-\frac{E_{a4}}{R} \cdot \left(\frac{1}{T(t)} - \frac{1}{T_{ref_4}}\right)\right) \cdot \left(\frac{n_A}{V}\right)^2$$

$$\vdots$$

The Reaction of Urethane

DAE model

- highly nonlinear Arrhenius kinetics
- 8 unknown parameters p

Measurements from different experiments with

- 3 measurement methods (A,C/D,E) with different variances
- different control functions u(t): temperature, feed 1, feed 2
- different control variables q: initial molar numbers, reaction volume

Structured nonlinear parameter estimation problem

A General Problem Formulation

Model: Ordinary Differential Equations (ODE)

$$\dot{y} = f(t, y, p, q, u)$$

y states

or: Differential Algebraic Equations (DAE)

$$\dot{y} = f(t, y, z, p, q, u)$$
$$0 = g(t, y, z, p, q, u)$$

y „differential“ states
z „algebraic“ states

p: (unknown) system parameters (PE) to determined
q: control parameters, u: control functions (given)

complex dynamics, instabilities, stiffness, discontinuities …

+ further constraints: initial or boundary condtions, positivity …

$$y(t_0) = y_0(p) \quad \text{or} \quad r(y(t_0), y(t_1), \cdots, y(T), p) = 0, \qquad p_i \geq 0$$

up to PDE (method of lines)

or: Optimal Control Problems

$$\min_{y,u} \quad \sum_j \gamma_j \Phi_M^j(y(T))$$

$$\text{s.t.} \quad \dot{y}(t) = f(t, y(t), u(t), p)$$

$$0 \le c(t, y(t), u(t), p)$$

$$0 = r(y(t_0), y(T), p)$$

y „differential" states
u control functions
γ weighting parameters
p system parameters

observed process is result of an optimization, e.g. in gait analysis

The Experimental Data

- measurements

$$\eta_{ij} = b_{ij}(t_i, y(t_i), z(t_i), p) + \varepsilon_{ij} \qquad j \in \mathrm{Ind}(i)$$

- measurement functions b_{ij}, with add'l calibration parameters
- measurement errors ε_{ij}
- from multiple experiments, under varying conditions
  - instationary states
  - stationary
  - bifurcations

each has specific model!

may comprise a priori information on parameters as “pseudo-measurements”

The Parameter Estimation Problem (DAE)

IWR Interdisciplinary Center for Scientific Computing

DAE process model

$$\dot{y} = f(t, y, z, p, q, u)$$
$$0 = g(t, y, z, p, q, u)$$
$$x := (y, z)$$

data

$$\eta_{ij} = b_{ij}(t_i, x(t_i), p, q) + \varepsilon_{ij}$$
$$\text{here: } \varepsilon_{ij} \in N(0, \sigma_{ij}^2)$$

determine p **and** x (parameters **and** states!)

$$\min_{x,p} \sum_{i,j} \frac{(\eta_{ij} - b_{ij}(t_i, x(t_i), p, q))^2}{\sigma_{ij}^2}$$
$$\dot{y} = f(t, y, z, p, q, u)$$
$$0 = g(t, y, z, p, q, u)$$
$$d(x(t_0), \ldots, x(t_f), p, q) = 0, \quad \text{or} \geq 0$$

Boundary value problem

The Multiple Experiment Case (DAE)

Interdisciplinary Center for Scientific Computing

DAE process model

$$\dot{y}_k = f_k(t, y_k, z_k, p, q_k, u_k)$$
$$0 = g(t, y_k, z_k, p, q_k, u_k)$$

data

$$\eta_{ijk} = b_{ijk}(t_{ik}, x(t_{ik}), p, q) + \varepsilon_{ijk}$$
$$\text{here}: \varepsilon_{ijk} \in N(0, \sigma_{ijk}^2)$$

determine p and x_k (k=1,...,# exp)

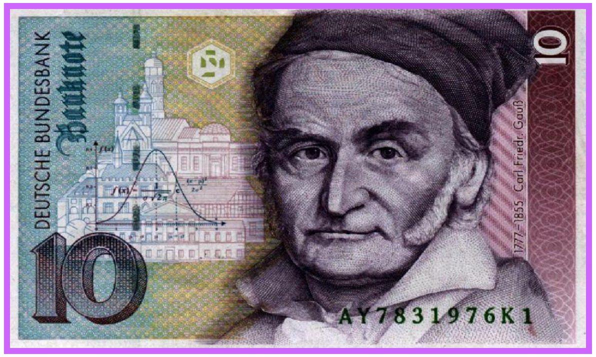
$$\min_{x_k, p} \sum_{k}^{\#\exp} \sum_{i,j} \frac{(\eta_{ijk} - b_{ijk}(t_{ik}, x_k(t_{ik}), p, q_k))^2}{\sigma_{ijk}^2}$$
$$\dot{y}_k = f_k(t, y_k, z_k, p, q_k, u_k)$$
$$0 = g_k(t, y_k, z_k, p, q_k, u_k)$$
$$d_k(x_k(t_{0k}), \ldots, x_k(t_{fk}), p, q_k) = 0, \quad \text{or} \geq 0$$

family of boundary value problems

The Choice of Norms

Interdisciplinary Center for Scientific Computing

- least squares norm: Legendre 1805, Gauss 1809 (normally distributed measurement error)



$$\|x\|_2^2 = \sum_i x_i^2$$

but much can be said in favour of

- ℓ_1-norm: Boscovic 1758, Laplace 1812
robust against outliers
(Laplace-distributed measurement error)

Kostina '02, '04

$$\|x\|_1 = \sum_i |x_i|$$

Direct Methods for Constrained Parameter Estimation

Direct "All-at-Once" Optimization
Boundary Value Problem (BVP) Methods

- the IVP approach: "single shooting"
  - integrate DAE over whole interval to yield $x(t;x_0,p)$ resp., solve OCP for given γ, p
  - eliminate all - infinite - variables in favour of unknown parameters γ, p,
  - plug into suitable optimizer

Direct "All-at-Once" Optimization
Boundary Value Problem (BVP) Methods

Interdisciplinary Center for Scientific Computing

- the IVP approach: "single shooting"
  - integrate DAE over whole interval to yield $x(t;x_0,p)$ resp., solve OCP for given γ, p
  - eliminate all - infinite - variables in favour of unknown parameters γ, p,
  - plug into suitable optimizer

- the BVP approach: discretize the DAE/OCP, and solve simultaneously
  - optimization problem
  - *discretized BVP as equality constraint or necessary conditions for discretized OCP*

  plus further constraints

in one loop

"all-at-once"

The Direct Multiple Shooting Method for Parameter Estimation in DAE

alternatives:
collocation on finite elements,
finite differences, ...
Biegler

parameterize/discretize DAE by the *multiple shooting* method, i. e.,

• choose suitable mesh

$$t_0 < t_1 < \ldots < t_m = t_f$$

• introduce state variables at nodes t_i

$$s_i^x \mathrel{\hat=} x(t_i), \;\; s_i^z \mathrel{\hat=} z(t_i)$$

as additional optimization variables

PARFIT

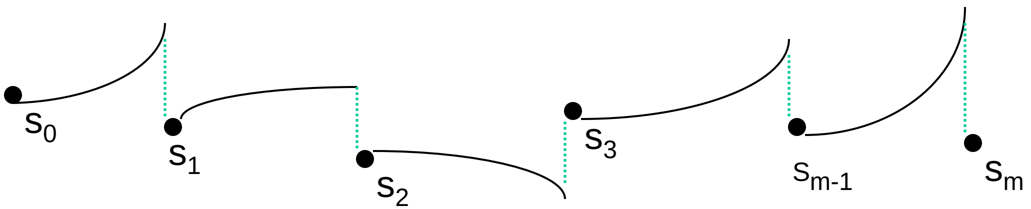
Bock, Bär, Schl. '78, '81, '83, '87 ff
Bock, Eich, Schl. '88, Kostina '01, '04, Kircheis ‘16

The Direct Multiple Shooting Method for Parameter Estimation in DAE

- integrate relaxed DAE on multiple shooting subintervals $[t_i, t_{i+1}]$

$$\dot{y} = f(t, y, z, p)$$
$$0 = g(t, y, z, p) - \alpha_i(t)\, g(t_i, s_i^y, s_i^z, p)$$

$$\begin{pmatrix} \alpha_i(t_i) = 1, \\ \alpha_i(t) \to 0 \quad (t \to \infty) \end{pmatrix}$$



$$y(t; s_i^y, s_i^z, p)$$

- jumps and relaxation terms must vanish at the solution → additional continuity and consistency conditions replace DAE

$$y(t_{i+1}; s_i^y, s_i^z, p) - s_{i+1}^y = 0$$
$$g(t_i, s_i^y, s_i^z, p) = 0$$

(i=0,...,m-1)

Result:
Constrained Nonlinear Least Squares Problem

[CNLS]

$$\min_X \|F_1(X)\|_2^2$$
$$\text{s.t. } F_2(X) = 0, \text{ or } \geq 0$$

$X = (p, s_0, s_1, \ldots, s_m)$
parameters and states

solution by Newton-type methods

X^0: initial guess!

$$X^{k+1} = X^k + t^k \Delta X^k$$

where ΔX^k solves a **constrained linear least squares problem (CLLS)**

[CLLS]

$$\min_{\Delta X} \|F_1(X^k) + J_1(X^k)\Delta X\|_2^2$$
$$\text{s.t. } F_2(X^k) + J_2(X^k)\Delta X = 0, \quad \text{or} \geq 0$$

$$J_i := \frac{\partial F_i}{\partial X}$$

solution by **generalized inverse** $\Delta X^k := -J^+(X^k)F(X^k)$ with $J^+ = J^+ J J^+$

Block-Sparse Structures of Jacobian

- super-structures from multiple experiments
- structure from multiple shooting
- sub-structures from spatial discretization of PDE
- sub-sub-structures from sparse state equations

$$
\begin{pmatrix}
\begin{smallmatrix} E_{L1} \\ U_{L1} \end{smallmatrix} & 0 & 0 & & 0 & \begin{smallmatrix} E_{G1} \\ U_{G1} \end{smallmatrix} \\
0 & \begin{smallmatrix} E_{L2} \\ U_{L2} \end{smallmatrix} & 0 & & 0 & \begin{smallmatrix} E_{G2} \\ U_{G2} \end{smallmatrix} \\
0 & 0 & \boxed{\begin{smallmatrix} E_{L3} \\ U_{L3} \end{smallmatrix}} & & 0 & \boxed{\begin{smallmatrix} E_{G3} \\ U_{G3} \end{smallmatrix}} \\
 & & & \ddots & & \vdots \\
0 & 0 & 0 & & \begin{smallmatrix} E_{LN} \\ U_{LN} \end{smallmatrix} & \begin{smallmatrix} E_{GN} \\ U_{GN} \end{smallmatrix}
\end{pmatrix}
$$

$$
J = \begin{bmatrix}
A_0^y & A_0^z & & \text{--- Algebraic + Invariant Cond.} \\
B_0^y & B_0^z & & \text{--- Initial + Interior Point Cond.} \\
G_0^y & G_0^z & -I & \text{--- Continuity} \\
 & & A_1^y \quad A_1^z & \\
 & & B_1^y \quad B_1^z & \\
 & & G_1^y \quad G_1^z \quad -I & \\
 & & \ddots & \\
 & & \ddots & \\
 & & G_{m-1}^y \quad G_{m-1}^z \quad -I & \\
 & & A_m^y \quad A_m^z & \\
 & & B_m^y \quad B_m^z & \\
D_0^y & D_0^z & D_1^y \quad D_1^z \quad \cdots \quad D_{m-1}^y \quad D_{m-1}^z \quad D_m^y \quad D_m^z & \\
E_0^y & E_0^z & E_1^y \quad E_1^z \quad \cdots \quad E_{m-1}^y \quad E_{m-1}^z \quad E_m^y \quad E_m^z &
\end{bmatrix};
\begin{bmatrix} a_0 \\ b_0 \\ m_1 \\ a_1 \\ b_1 \\ m_2 \\ \vdots \\ \vdots \\ m_m \\ a_m \\ b_m \\ d \\ e \end{bmatrix} = -F
$$

\# experiments 1 - 100

\# multiple shooting points typically10 - 40

must exploit special block structures $\longrightarrow$ condensing

FAQ: Why Multiple Shooting?

- key property: discretized states as add'l optimization variables
  - allows ***better initial guesses***, using information about process, which helps avoid "far away" local minima
  - ***reduces nonlinearity***, even down to one-step convergence
  - is ***numerically stable***, suitable even for highly unstable, e.g. chaotic dynamics
- efficient parallel implementation
- adaptive accuracy discretization strategies
- add'l advantage of multiple shooting
  - state-of-the-art solvers for DAE applicable
  - treatment of discontinuities (hybrid systems) e.g. phase changes, hysteresis, ...

"adaptive accuracy" realized through integrator

An Unstable Test Problem

Unstable Test Problem

- state equations:

$$\dot{x}_1 = x_2 \qquad , x_1(0) = 0$$
$$\dot{x}_2 = \mu^2 x_1 - (\mu^2 + p^2) \sin pt \qquad , x_2(0) = \pi$$

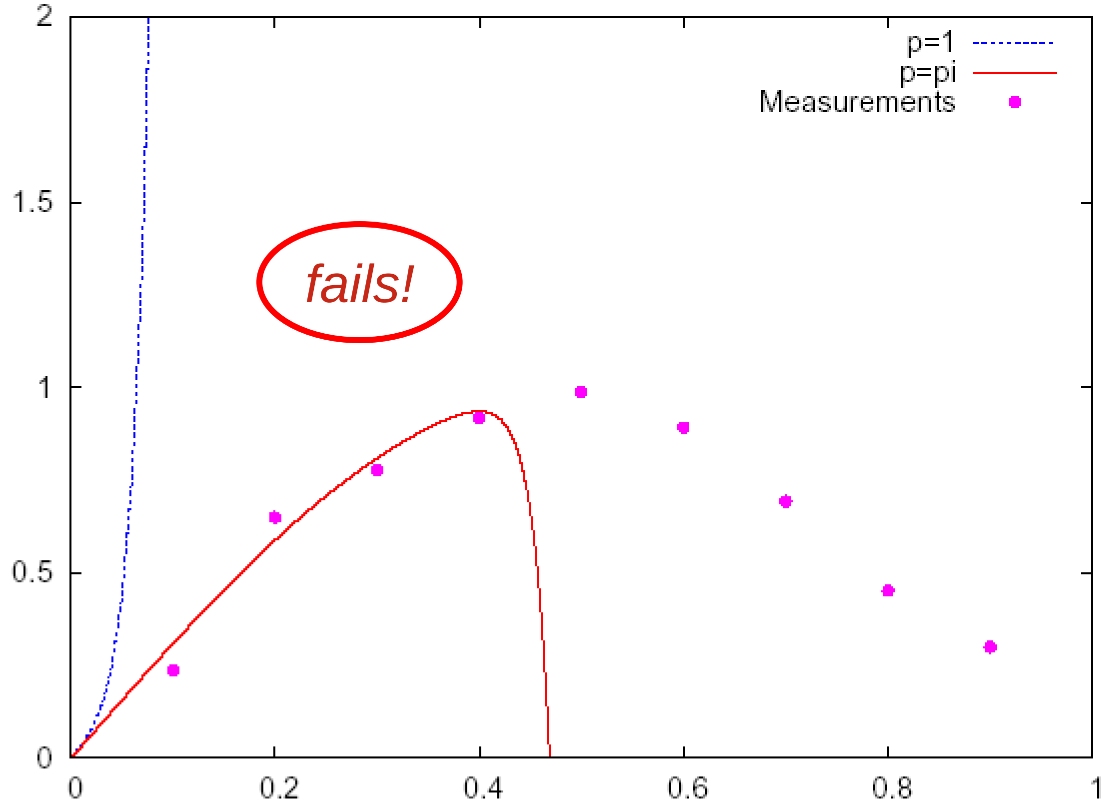
$t \in [0,1]$

- special solution for "true" parameter value $p = \pi$:

$$x_1(t) = \sin \pi t$$
$$x_2(t) = \pi \cos \pi t$$

- pseudo random measurement noise, $\sigma = 0.05$

Unstable Test Problem - Single Shooting



initial trajectory with p=1, and with p=float(π) in 64 bit

initial value problem is extremely ill-conditioned!

Unstable Test Problem - General Solution

$$x_1 = \sin pt + \varepsilon_1 \sinh \mu t + \varepsilon_2 \cosh \mu t \,; \qquad \varepsilon_1 = \frac{x_2(0) - p}{\mu}$$

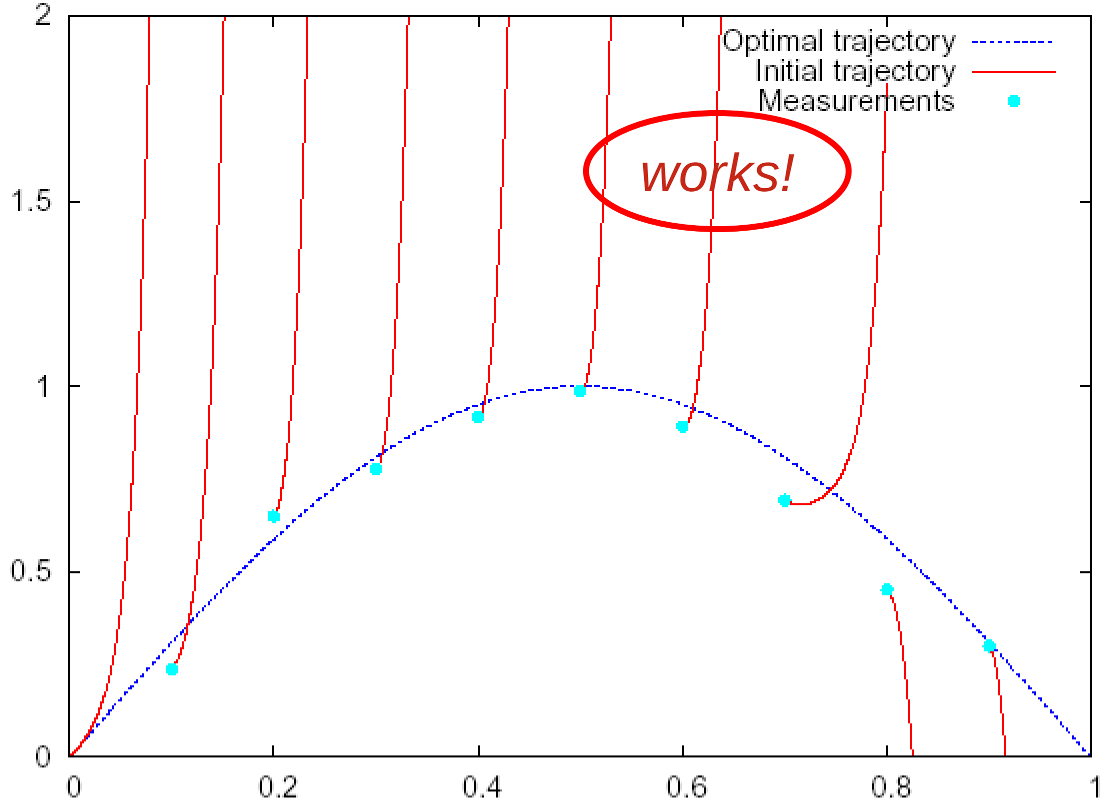
$$x_2 = p \cos pt + \varepsilon_1 \mu \cosh \mu t + \varepsilon_2 \mu \sinh \mu t; \qquad \varepsilon_2 = x_1(0)$$

Eigenvalues of $f_x(t, x(t), p)$ are $\lambda_{1,2} = \pm \mu$

Error propagation: exp(+/- μt)!

μ =60, i.e. error propogation over [0,1] is 10^{27} - highly unstable!

Unstable Test Problem - Multiple Shooting



initial trajectory for p=1 - convergence after 4 iterations

parameter estimation problem is well-conditioned!

Efficiency of Boundary Value Problem Methods

Theorem

Assumptions

Dense exact data for all states available

Model equations linear in parameters

Initial guesses for states: given data

Length of multiple shooting (resp. collocation) intervals $h \to 0$

Then

One step convergence to true parameter value p^*

$$p^1 = p^0 + \Delta p^0 = p^* + O(h)$$

Reduction of Nonlinearity by Decoupling

Advantages of BVP approach also in case of non-dense noisy data

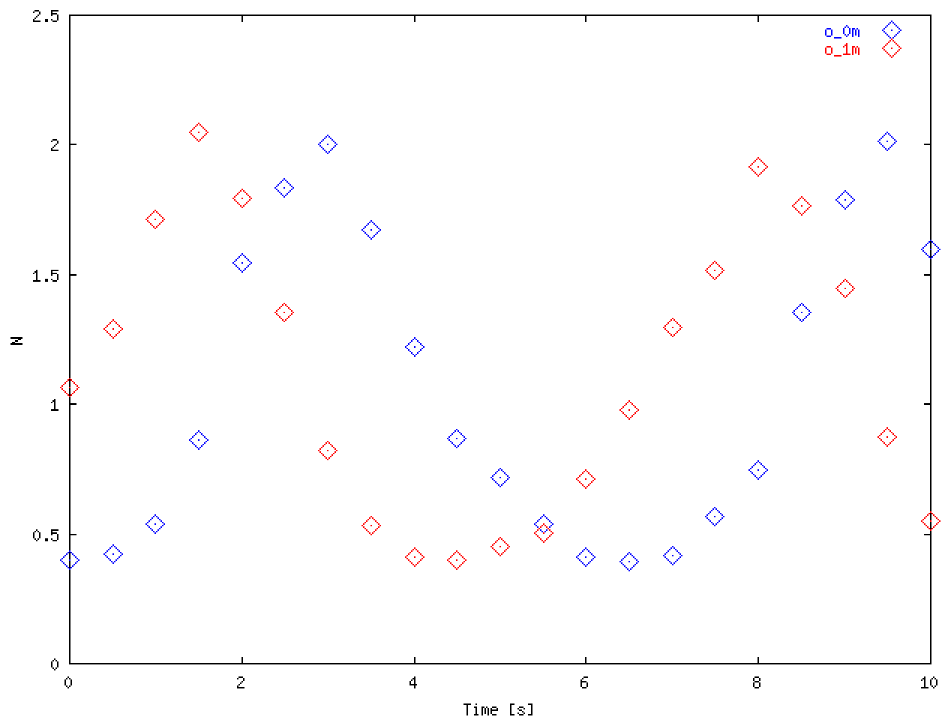
Lotka-Volterra Problem

Lotka-Volterra: Model and Data

$$\dot{x}_1 = -p_1 x_1 + p_2 x_1 x_2$$
$$\dot{x}_2 = +p_3 x_2 - p_4 x_1 x_2$$

x_1: predators
x_2: preys

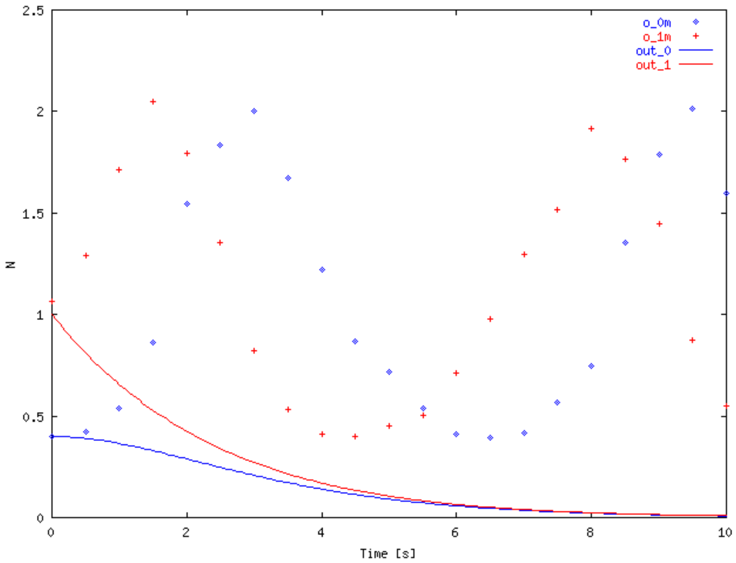
DE linear in par‘s



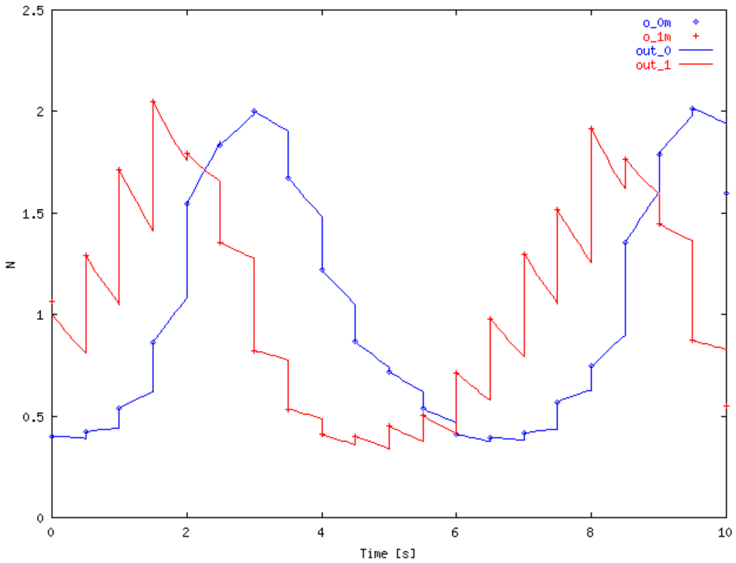
Data: $\sigma = 5\%$

Comparison: Single vs. Multiple Shooting

Single Shooting



Multiple Shooting



Convergence
after 20 iterations

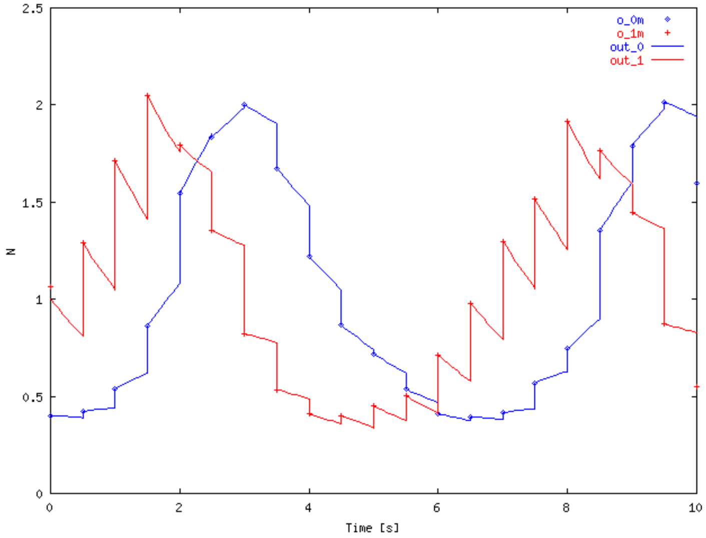
Convergence
after 4 iterations

Initial guesses $p_1 = 0.5$ $p_2 = 0.5$

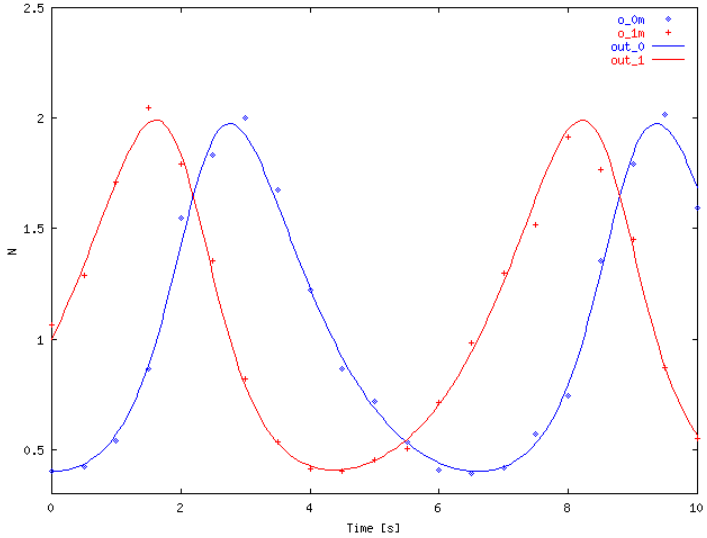
$p_3 = -0.5$ $p_4 = -0.2$

Lotka-Volterra: Solution with Multiple Shooting

Initial Trajectory



Solution Trajectory



Initial guesses:

$p_1 = 0.5$ $p_2 = 0.5$

$p_3 = -0.5$ $p_4 = -0.2$

Solution:

$p_1 = 1.01 \pm 0.02$ $p_2 = 1.01 \pm 0.03$

$p_3 = 0.99 \pm 0.02$ $p_4 = 1.01 \pm 0.03$

Some Algorithmic Features

Evaluation of CLLS
Computation of 1st and Higher Order Derivatives

the crucial requirement for practical use:
numerics must be "derivative-free" *for the user!*

- adaptive integrators for ODE and relaxed DAE
- fast and error controlled computation of 1st and higher order derivatives
  - combining "automatic differentiation" of model equations and
  - "internal differentiation" of adaptive discretization scheme
  - treatment of implicitly given discontinuities and jumps in dynamics
  - in forward or *reverse (adjoint)* mode

e.g. **DAESOL, RKFSWT**

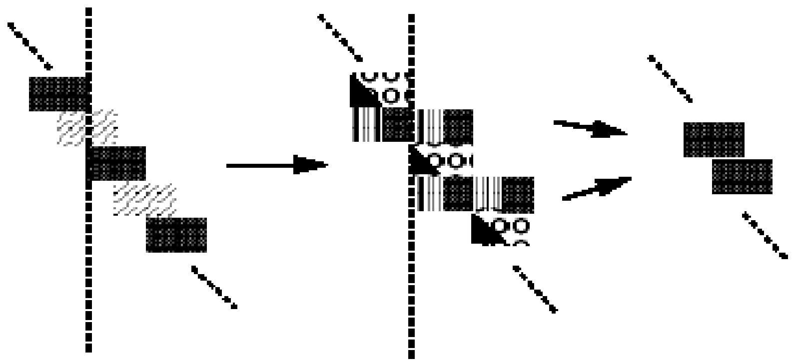
Bauer et al. '98
Albersmeyer '05
Kirches '06

Parallel Evaluation and Decomposition

1. Evaluation of functions and gradients: parallel on interval level
2. Parallel condensing

log(m) - algorithm

treatment of
staircase system



$[G_i^l | G_i^r] \longrightarrow [R_i | 0] \cdot T_i$

Variants: Orthogonal transformations for unstable systems
Block Gauss elimination for stable systems

Fast Sequential Solution
Reduced Generalized Gauss Newton

Idea:

Use initial, multipoint, DAE consistency conditions, ...

to reduce number of directional derivatives of IVP solutions to minimum

1. DAE consistency $A^y \Delta S^y + A^z \Delta S^z + \alpha = 0$

   Computation of solution mf $N\delta + \Delta\hat{S},\ \delta$ free

2. Continuity $G^y \Delta S^y + G^z \Delta S^z - \Delta S_+^y + \mathrm{m} = 0$

   $(\#\,\mathrm{Gradients}: \#\,\Delta S^y + \#\,\Delta S^z)$

   Insertion of solution mf $\left[G^y\ G^z\right] N\delta + \left[G^y\ G^z\right]\Delta\hat{S} - \Delta \mathrm{S}_+^{\mathrm{y}} + \mathrm{m} = 0$

   $(\#\,\mathrm{Gradients}: \#\,\delta + 1)$

\# Needed directional derivatives = # Degrees of freedom + 1 → **PDE**

Treatment of Condensed System

Large-scale linear constrained system is reduced to condensed system in n variables

$$
\begin{aligned}
\min \quad & ||A_1 v + b_1||_\alpha \\
\text{s.t.} \quad & A_{2E} v + b_{2E} = 0 \\
& A_{2I} v + b_{2I} \geq 0 \\
& v \in R^n \\
& \alpha = 1, 2, \infty
\end{aligned}
$$

$n \leq$ number of states + parameters   n small !

Solution: l_1, l_∞ Modifications of „Adaptive Method“ (Gabasov, Kirillova, Kostina, ...

l_2 Orthogonal Transformations & Active Set Strategies and Elimination

Detection of ill-posedness, rank deficiency → regularisation strategies

Regularisation Strategies

Equality constrained **underdetermined** linear system

- Determine rank n_r ($n_r < n$) of linear system, e.g. in course of decomposition
- Determine solution manifold $Cx + d = 0$ (C full rank n_r, $x \in \Re^n$)
- Solve minimum norm problem

$$\min_x \quad \|x\|_\alpha$$
$$\text{s.t.} \quad Cx + d = 0$$
$$\alpha = 1,2,\ldots$$

- Constrained linear least squares or linear programming problem, resp.
- Note: In l_1 case 0-components can be expected!

Convergence of Constrained Gauss-Newton

- active set constant near solution → equality constrained problem

  convergence proof based on 2 principal assumptions:

$$\left\|(J^{+}(X-\Delta X)(J(X-t\Delta X))-J(X))(\Delta X)\right\| \le \omega t \left\|\Delta X\right\|^{2} \quad \text{with } \omega < \infty$$

$$\left\|(J^{+}(Z)-J^{+}(X))R(X)\right\| \le \kappa(X)\left\|Z-X\right\| \le \kappa\left\|Z-X\right\| \quad \text{with } \kappa < 1$$

$$\forall\, t \in [0,1],\quad \forall\, X, Z \in D \quad \text{with } \Delta X = J^{+}(X)F(X),\quad R(X) = F(X) - J(X)\Delta X$$

- **local *linear* convergence** of full step method with asymptotic rate κ
- **advantage:** method not attracted by large residual local minima X^{*},

  so called "statistically unstable" minima - cannot be interpreted as continuous deformation of "true parameter" !
- **global convergence:** by efficient new strategies based on "affine invariance" principles - guarantee full step in local domain of convergence

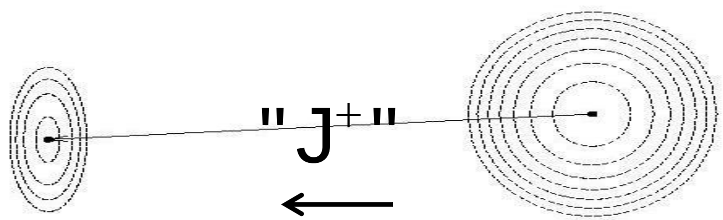
Bock, Kostina, Schl. '00

Statistical Assessment of Solution

Statistical Sensitivity Analysis for Constrained Case

- need to know uncertainty of parameter estimate $X^*(\varepsilon)$ depending on measurements errors $\varepsilon \in N(0,\beta^2 I)$

$X^*(\varepsilon) \,\dot{\in}\, N(X^*, C)$



$\varepsilon \in N(0,\beta^2 I)$

- first order expansion:

$$X^*(\varepsilon) - X^* \doteq -J(X^*)^+ \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}, X^* := X^*(0)$$

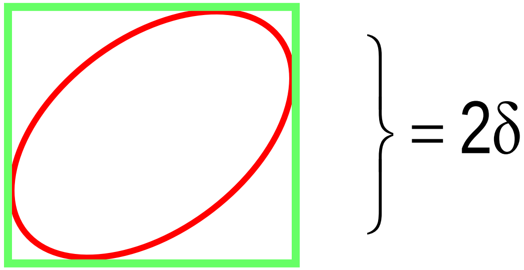
- covariance matrix approximation:

$$C := E\left(J(X^*)^+ \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}^T J(X^*)^{+T} \right) = J(X^*)^+ \begin{pmatrix} \beta^2 I & 0 \\ 0 & 0 \end{pmatrix} J(X^*)^{+T}$$

Statistical Sensitivity Analysis for Constrained Case

- confidence ellipsoid G, includes "true value" with error probability $\approx \alpha$

$$G := \{X \mid X - X^* = -J^+(X^*)\begin{pmatrix}\varepsilon \\ 0\end{pmatrix}, \|\varepsilon\|_2^2 \leq \gamma(\alpha)\}, \gamma(\alpha) := v_1 F^{1-\alpha}(v_1, v_2)$$



- Lemma: G can be enclosed by confidence box

$$\prod [X_i^* - \delta_i, X_i^* + \delta_i], \quad \delta_i = C_{ii}^{1/2}\gamma(\alpha))^{1/2}$$

- need to compute only $C_{ii}^{1/2}$, "standard deviations" of parameters

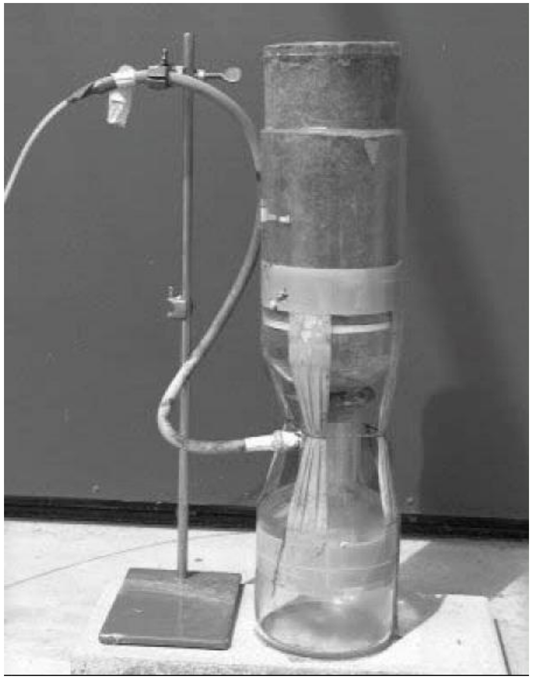
→basis for optimum experimental design

Applications

Transport and Degradation of Xenobiotics in Soil

in cooperation with BASF The Chemical Company

Transport and Degradation of Xenobiotics in Soil



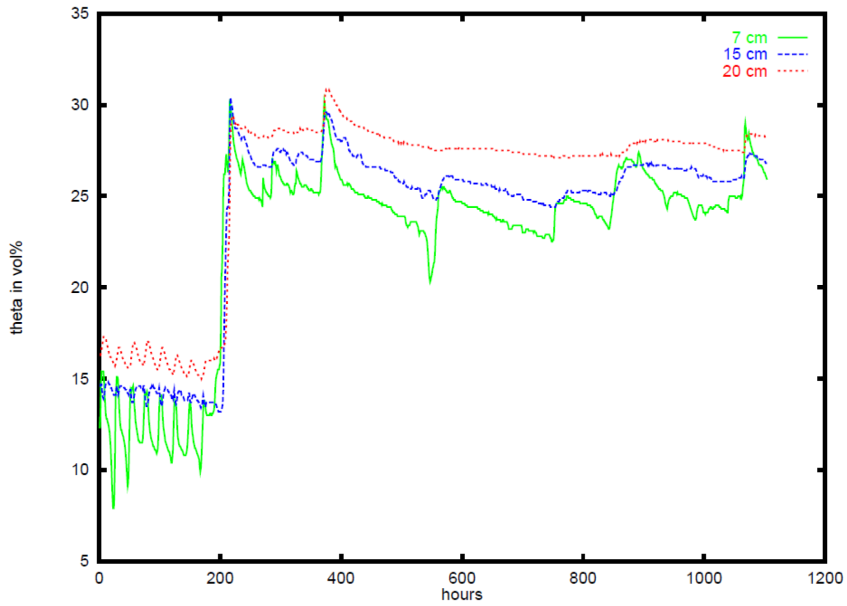
mini-lysimeter

- Investigation of fate of xenobiotics
- Expensive lysimeter experiments for registration
- To be replaced by computer experiments
- **Here: parameter estimation**
- **Later: optimal mini-lysimeter experiments**
  - **optimal irrigation**
  - **optimal solute application**
  - **Optimal sampling**

Altmann-Dieses, Bock, Schl., Richter '02

Field experiment: Water Transport (K. Aden)

- loamy sand without vegetation
- time-domain reflectrometry (TDR): hourly readout
- measurements of water content θ in 7, 15 and 20 cm
- period: Oct 28, 1997 - Dec 13, 1997



PDE-Model: Richards Equation

$$\frac{\partial\theta}{\partial t} = \frac{\partial}{\partial z}\left(D(\theta)\frac{\partial\theta}{\partial z} - K(\theta)\right)$$

$$K(\theta) = K_s\Theta^{1/2}\left[1 - \left(1 - \Theta^{n/(n-1)}\right)^{1-1/n}\right]^2, \qquad \Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r}$$

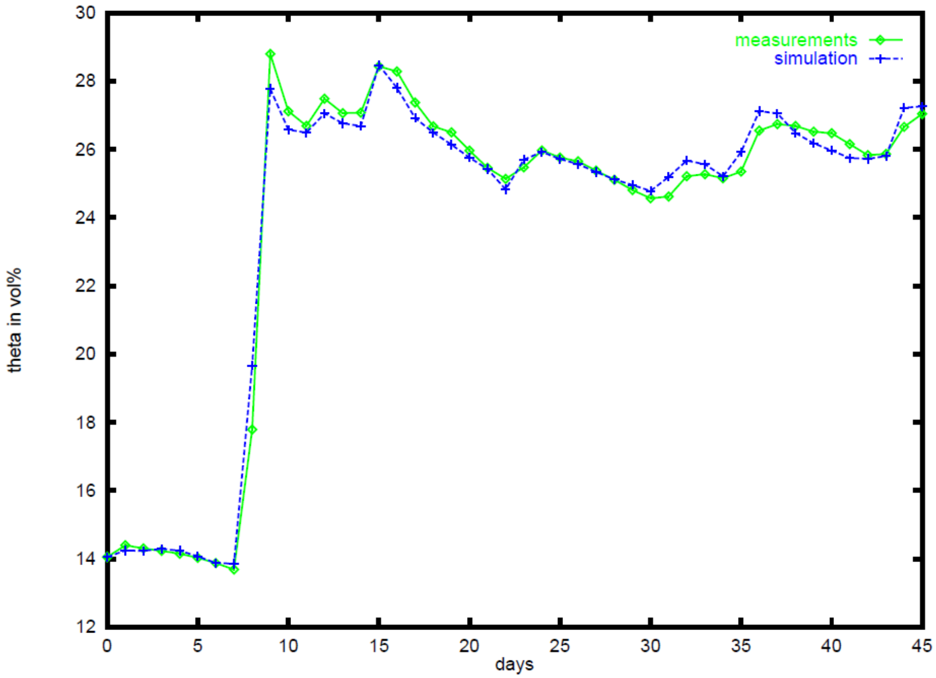
$$D(\theta) = K(\theta)\bar{C}(\theta)$$

$$\bar{C}(\theta) = \frac{1}{\alpha nm}\left(\Theta^{-1/m} - 1\right)^{-m}\Theta^{-1/m}\frac{1}{\theta - \theta_r}, \qquad m = 1 - \frac{1}{n}$$

- Initial condition: Linear interpolation of $\theta_{7\text{cm}}$, $\theta_{15\text{cm}}$, $\theta_{20\text{cm}}$ at the start of experiments (Oct 28, 1997)
- Upper boundary: Dirichlet condition (TDR data in 7 cm)
- Lower boundary: Dirichlet condition (TDR data in 20 cm)

Transport and Degradation of Xenobiotics in Soil

Result: Estimates for n, α and K_s



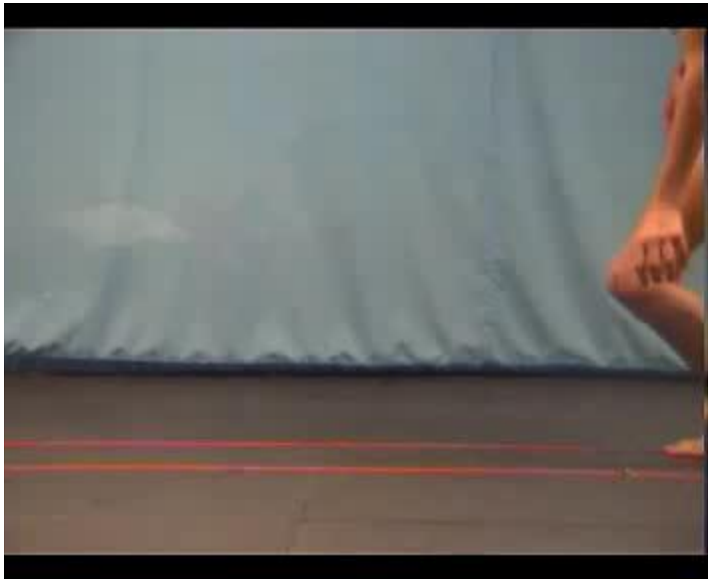
	guess	estimate
n	1.5	1.262 ± 0.0024
α	0.05	0.0324 ± 0.0024
K_s	35.0	20.92 ± 1.68

	α	K_s
n	0.14	-0.61
α	-	-0.94

Identification of Cerebral Palsy Gaits

Cerebral Palsy Gaits

before surgery



after surgery

Assumption: Movement is optimal

Task: Find suitable optimal control problem!

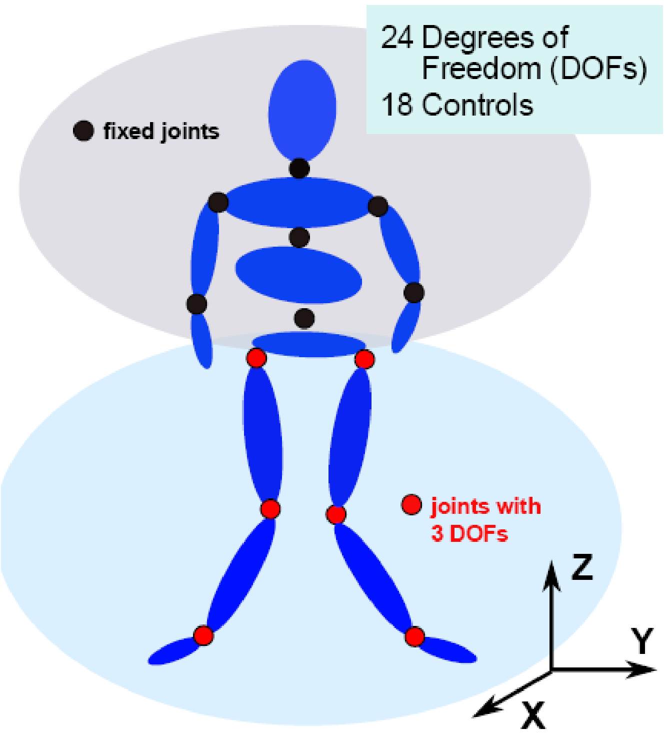
... K. Hatz in coop with S. Wolf (Orthopedics HD)

Identification of Cerebral Palsy Gaits

Model for patient‘s motion:

- 48 states q: 3 global coordinates, 3 global angles, 18 local joint angles (generalized coordinates), and the corresponding velocities $\dot{q}$
- 18 controls u
- constrained multibody system
- formulated with HuMAnS Toolbox (INRIA, France)

Multibody System Dynamics



patient data from Orthopedics Lab, resp. Deleva data

Inverse Optimal Control for Cerebral Palsy Gaits

$$\min_{y,u,p,\gamma} \sum_{i,j} \frac{(\eta_{ij} - b_{ij}(t_i, y(t_i))^2}{\sigma_{ij}^2}$$

s.t.

$$\min_{y:=(q,\dot{q}),\, u} \sum_k \gamma_k C_k[q, \dot{q}, u, p]$$

Optimal control model

s.t. $\dot{y}(t) = f(t, y(t), u(t), p)$ — MBS dynamics

$0 \le c(t, y(t), u(t), p)$ — control/path constraints

$0 = r(y(t_0), y(T))$ — boundary conditions

$$\sum_k \gamma_k = 1, \quad \gamma_k \ge 0 \;\; \forall k$$

Possible criteria C_k: stability, energy, duration,...

Challenges: discontinuities and jumps in states, large-scale, ...

Efficient Direct All-at-Once Approach

after discretization of optimal control problem with direct multiple shooting

$$\min_{x,p,\gamma,\lambda,\mu} \sum_{i,j} \frac{(\eta_{ij} - b_{ij}(t_i, y(t_i))^2}{\sigma_{ij}^2}$$

s.t.

$$0 = y(t_{i+1}; t_i, s_i, w_{i,}p) - s_{i+1}$$

$$0 \leq \tilde{c}(s_0, \dots, s_m, w, p)$$

$$0 = r(s_o, s_m)$$

Multiple Shooting discretization

x:=(s,w)
s multiple shooting variables
w control variables

$$0 = \nabla_x L(x, \gamma, p, \lambda, \mu)$$

$$0 \leq \mu$$

$$0 \geq \mu^T \tilde{c}(s_0, \dots, s_m, w, p)$$

KKT conditions

L Lagrangian
λ,μ adjoints

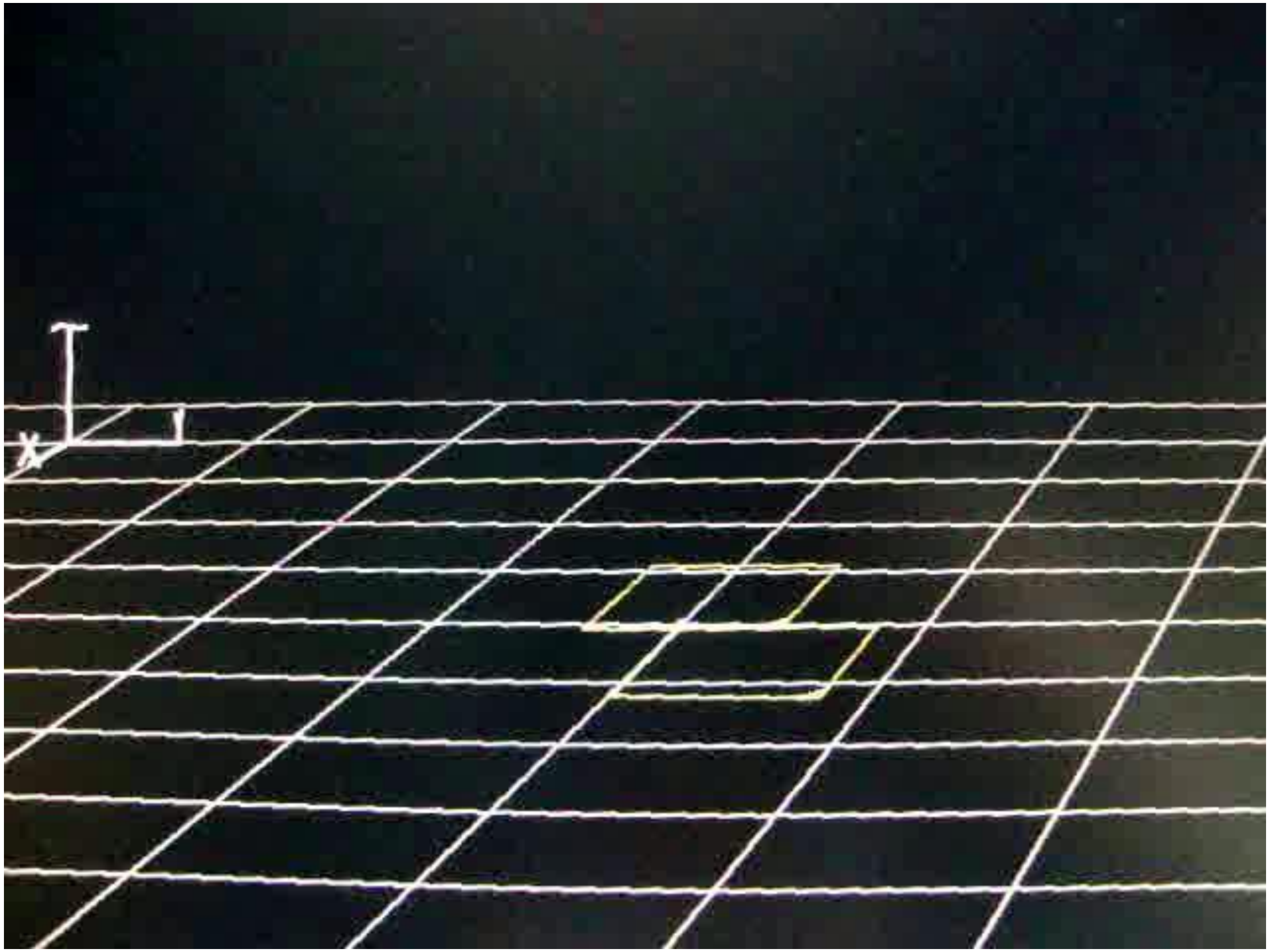
$$1 = \sum_k \gamma_k$$

$$0 \leq \gamma_k \qquad \forall k$$

objective regularization

large-scale constrained ls-problem with complementarity condition

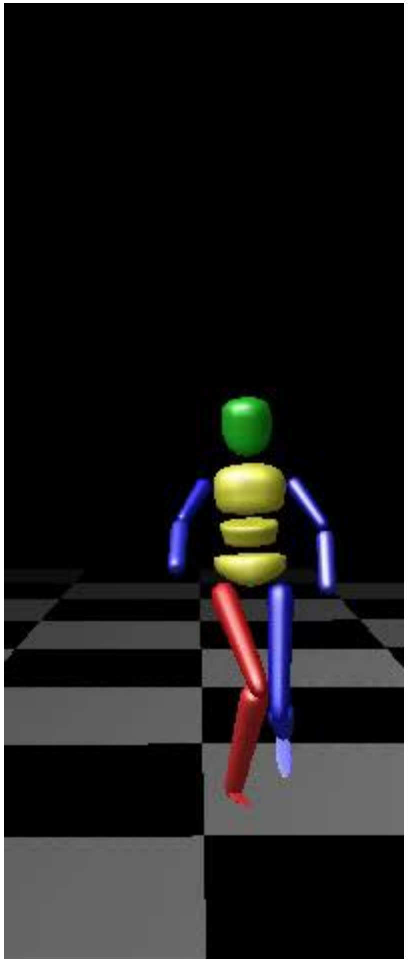
Measurements for Cerebral Palsy Gait



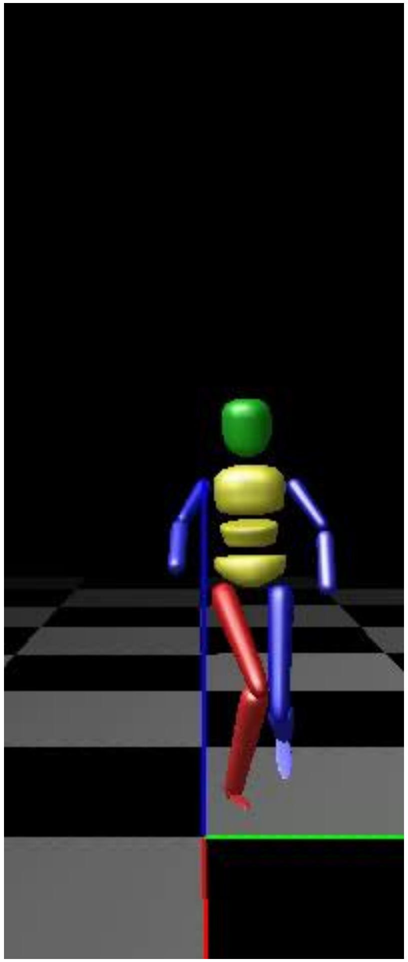
Vicon data

Identified Cerebral Palsy Gait

measured gait

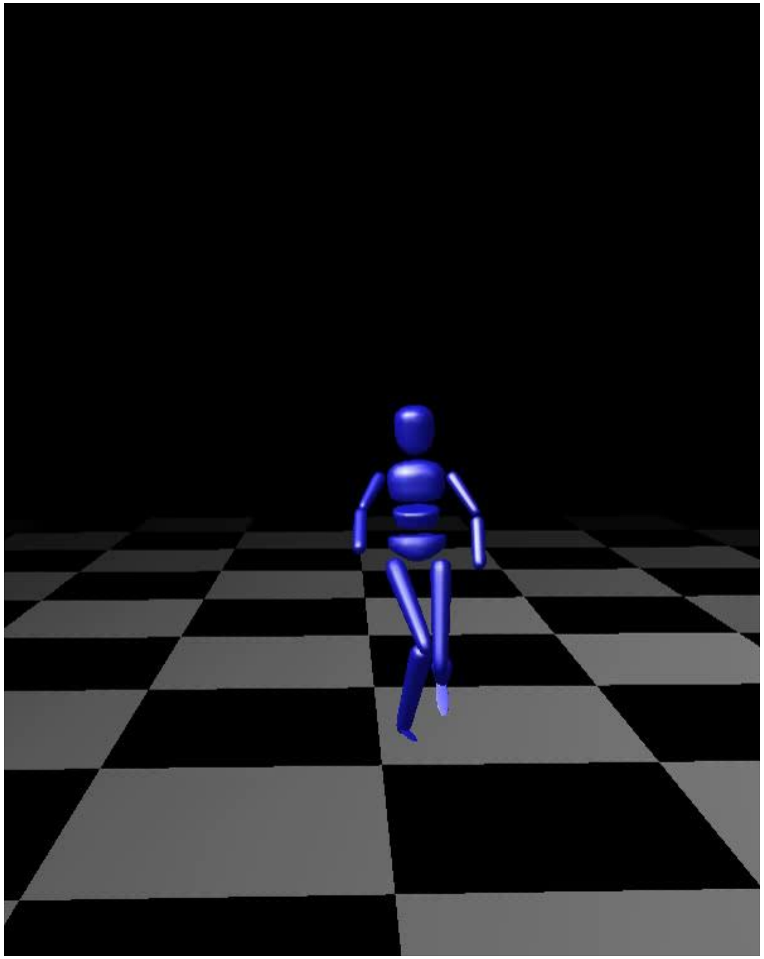


identified gait



very good agreement

Measured and Estimated CP Gait



... K. Hatz

green: measured gait   blue: estimated optimal OCP gait

Summary

- Parameter Estimation in Differential Equations
- Optimization Boundary Value Problems
  Hierarchical Optimization Problems
- Direct Multiple Shooting
- Some Applications

Ongoing work

Numerical Tools for Inverse Optimal Control
Nonlinear Optimum Experimental Design
Dual Control: Experimental Design in Control

Interdisciplinary Center for Scientific Computing

Thank you very much for your attention!