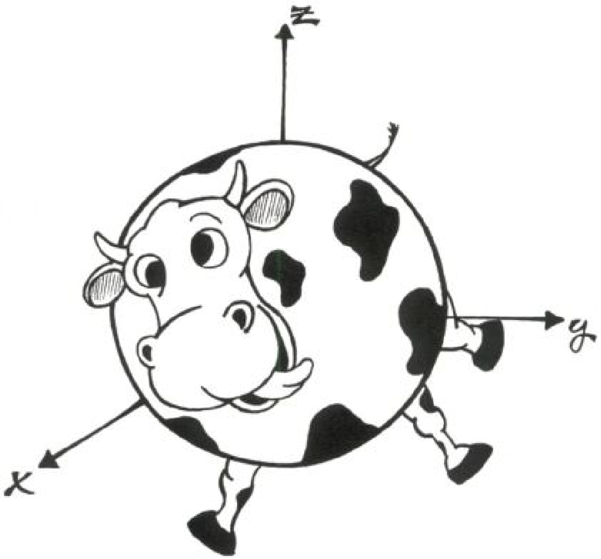
"Given a spherical cow"

Challenges that Engineering Problems Provide for Mathematical Research

Hans Georg Bock
Interdisciplinary Center for Scientific Computing (IWR)
Ruprecht-Karls-Universität Heidelberg



Mathematics for Key Technologies and Innovation
Warsaw, February 21-22, 2008

Overview

• detailed mathematical models as a key for understanding and mastering complex – engineering – problems

• intelligent algorithms for sophisticated optimization and inverse simulation rather than forward simulation

• optimization of dynamical processes: many real world engineering problems looking for new mathematical solutions

  • hybrid models: discontinuous dynamics and mixed integer optimization and optimal control

  • requirements of real-time optimization and control

  • experimental design for non-linear dynamical processes, uncertainties in models

• applications from Joint ICM/IWR Graduate College and the new "Heidelberg Graduate School Mathematical and Computational Methods for the Sciences"

Modeling, Simulation and Optimization of Engineering Processes

Dynamical Processes

dynamical processes

in many areas ....

- robots working at an assembly line
- operating a distillation column
- flying a power generating kite
- driving an automobile
- controlling a chemical reactor
- ....

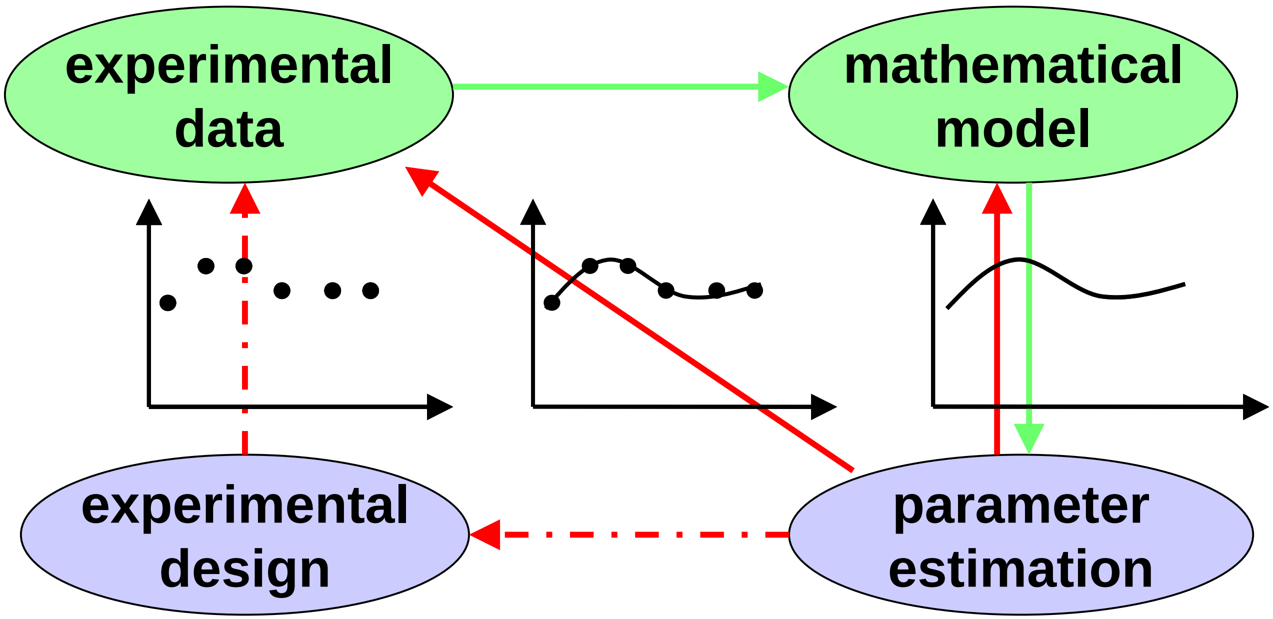
How to choose the decision variables such that the system or the process operates optimally?

Demand for Optimization is Ubiquitous

- increase productivity
- improve product quality
- conserve resources
- reduce waste
- increase safety
- reduce experimental costs
- ....

Important: strict consideration of all kinds of side conditions!

Model-Based Optimization Problems



Aims: scientific insight into process → simulation of scenarios → process optimization and control

The Model Class: (Partial) Differential Algebraic Equations

$$
\begin{aligned}
B(x,z,p)\,\dot{x} &= f(x,z,u,p) \\
0 &= g(x,z,u,p)
\end{aligned}
$$

x „differential“ states
z „algebraic“ states

u: control functions, p: system parameters

- ordinary differential equations
- mechanical DAE

$\vdots$

- reactive flow problems

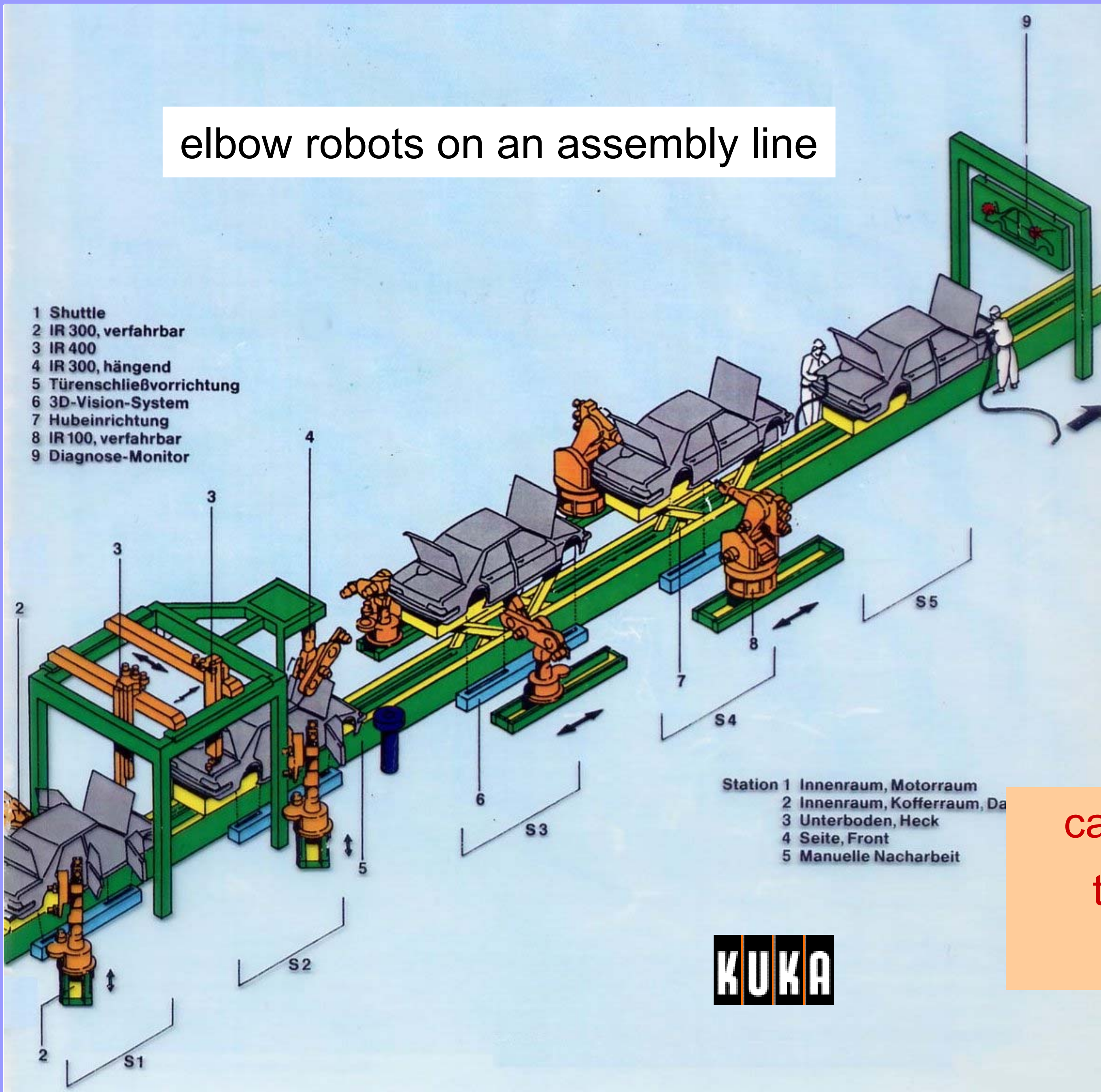
typical properties

- nonlinear
- hybrid: state dependent discontinuities
- different time scales

Given an initial state $x(t_0)$, $z(t_0)$, choose control functions u and parameters p in order to minimize some cost function $\Phi[x,z,u,p]$ subject to state and control constraints $\Psi[x,z,u,p] \geq 0$

Example: Time Optimal Robot Motion

elbow robots on an assembly line



- welding
- assembling
- gluing
- transporting

can we speed up the bottleneck maneuvers?

Time Optimal Motion of Robots

Traditional approach:

- move along prescribed simple geometric curves
- try to avoid dynamic interactions

Time optimal solution:

- make use of nonlinear interactions, i.e.,
- exploit gravity to help start and stop motion
- reduce or increase inertia and
- produce centrifugal forces to accelerate or decelerate rotational motion

time optimal solutions are up to 3 x faster!

Complex Case: Descriptor Form Models, Index 3 DAE

e.g. closed kinematic loops - changing DoF due to contact - natural coordinate modeling

- index 3 Differential Algebraic Equations (DAE)

$$
\begin{aligned}
&M(p)\cdot\ddot{p} - G(p)^T\lambda = B(p,\dot{p}) + D\cdot u(t), G(p) = \nabla g(p)^T \\
&g(p) = 0 \qquad (*)
\end{aligned}
$$

- (*) reducible to index 1 DAE (by 2 differentiations)

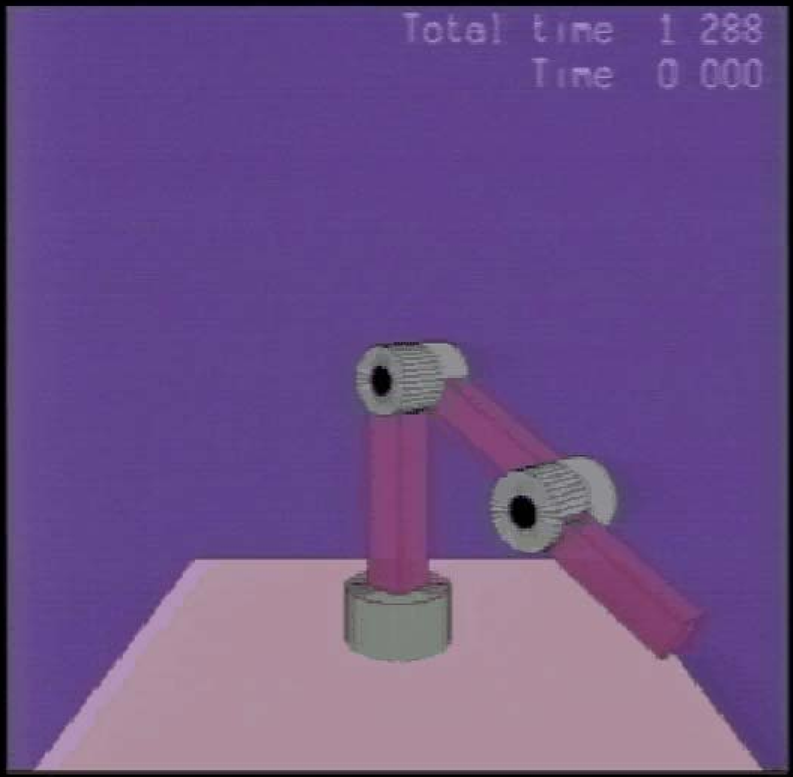
$$
\begin{aligned}
&M(p)\cdot\ddot{p} - G(p)^T\lambda = B(p,\dot{p}) + D\cdot u(t), G(p) = \nabla g(p)^T \\
&G(p)\cdot\ddot{p} = \gamma(p,\dot{p})
\end{aligned}
$$

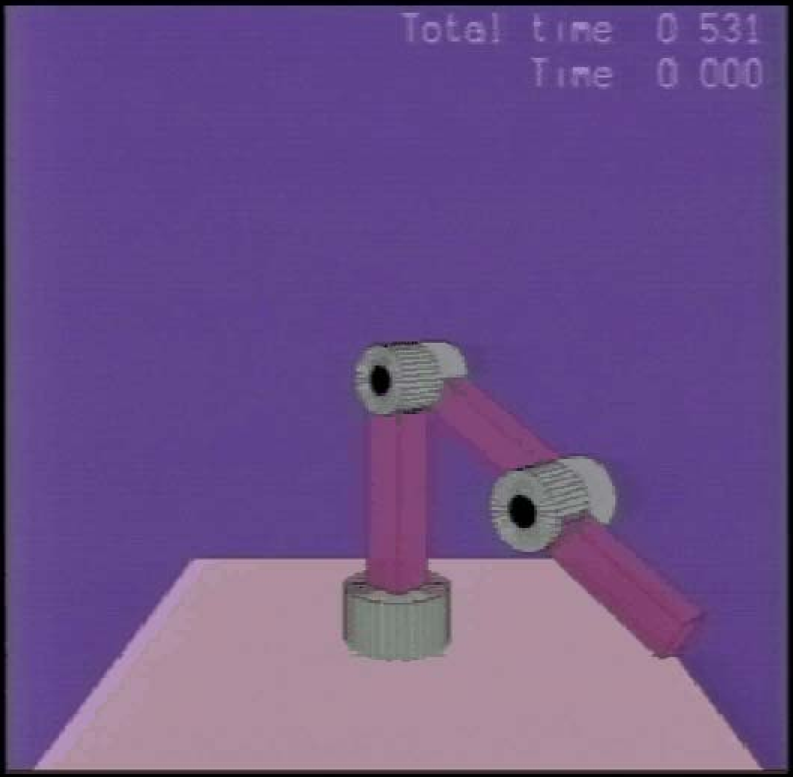
with invariants

$$
g(p) = 0, \qquad G(p)\cdot\dot{p} = 0
$$

traditional approach

time optimal solution
more than twice as fast





(slow motion x5)

Hybrid Models I:

Discontinuous Dynamics

Discontinuous Dynamics

Mathematical problem:

- right hand sides of DE (source terms) have state-dependent discontinuities, e.g.,

$$
\begin{aligned}
B(x,z,p;i)\,\dot{x} &= f(x,z,u,p;i) \\
0 &= g(x,z,u,p;i) \\
i &= h(x,z,u,p)
\end{aligned}
$$

- **i is integer variable! x, z, u, p may change spaces if i changes**

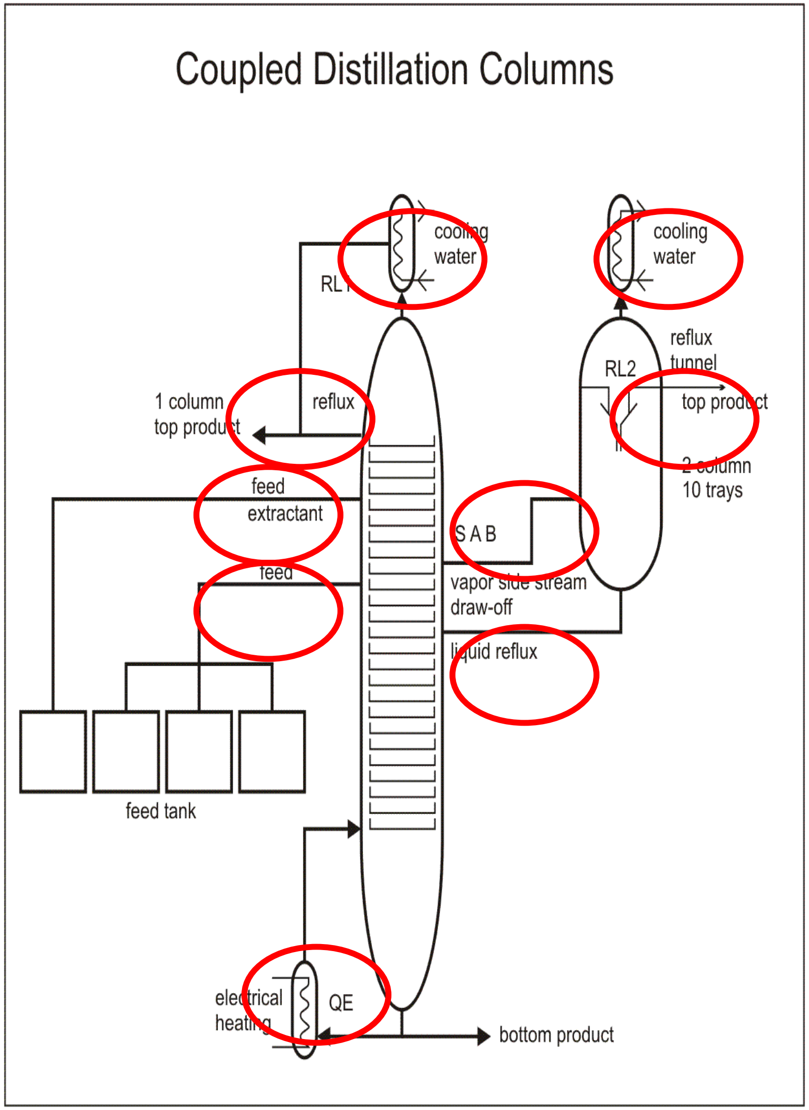
- origin: multi-scale processes, discontinuous modeling of fast time scales
- physical reasons: e.g. automatic electrical switches or transmissions, opening or closing of valves, multiple phases (gas-liquid), mechanical contact events, closing kinematic loops, etc.

Example: Control of a Distillation Column

Distillation columns are "energy guzzlers"!

- 7% of total energy in the US
- 70-80% of energy consumption at BASF site

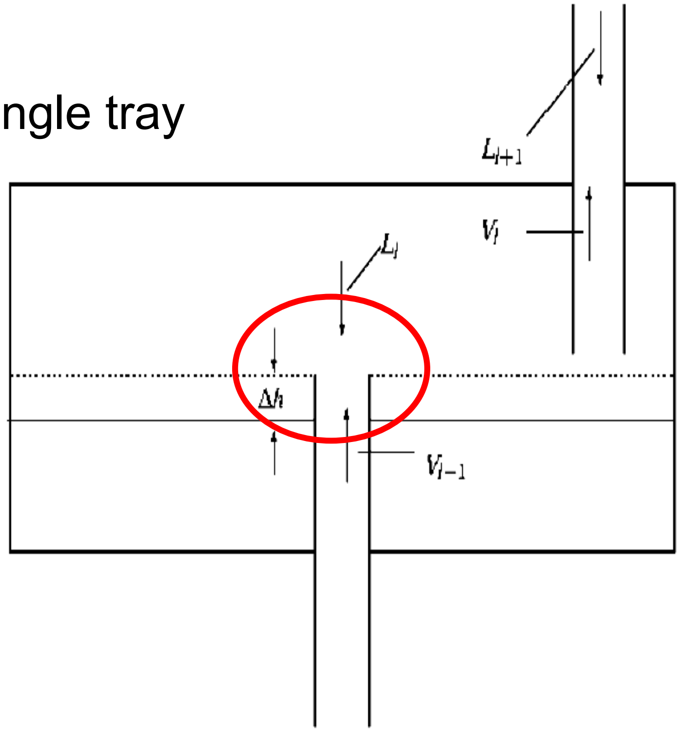
Control of a Distillation Column: Problem



- separate ternary mixture by conti-process
- minimize energy during start-up phase – transition to steady state operation
- principal operation:
- fill column with inert gas, heat, then add mixture
- controls: feed streams, reflux ratios, heating, cooling

Control of a Distillation Column: Model

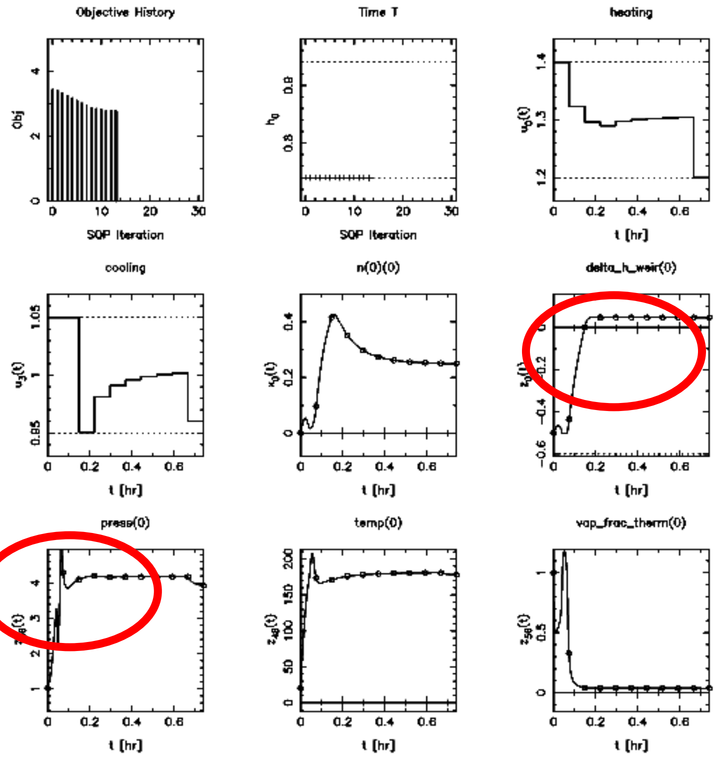
single tray



detailed physical model of gas liquid mixture

- temperature and pressure dependent phase transitions
- Antoine kinetics, phase specific densities
- non-differentiable dynamics at weir overflow
- 1012 state variables, a priori unknown switching structures

Control of a Distillation Column: Result



- constraints on purities, pressures, temperatures,
- energy reduced by 30 % compared to heuristic operation (constant heating)
- start-up phase reduced from 45 to 30 minutes
- over 300 discontinuities: phase changes and weir overflows

Brandt-Pollmann '04

Hybrid Models II:

Mixed-Integer Optimization and Optimal Control

Mixed Integer Optimal Control

Mathematical problem:

- in addition, **controls u(t) are functions with integer values!**

$$\begin{aligned} B(x,z,p;i)\,\dot{x} &= f(x,z,u,p;i) \\ 0 &= g(x,z,u,p;i) \\ i &= h(x,z,u,p) \end{aligned}$$

- origin: multi-scale processes

- physical reasons: e.g. optimal control of **gears** or on-off **switches**, opening or closing of **valves**, etc. modeled as **discrete decisions**

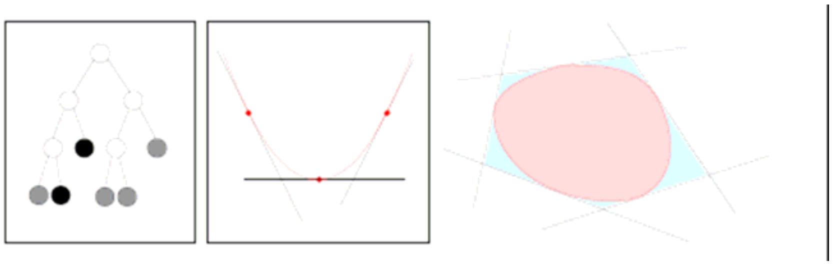
Mixed Integer Optimal Control

Principal solution methods:

Direct methods:

- piecewise constant discretization of control functions
- on time grid $t_0 < t_1 < ... < t_N = T$

Classical approach (MINLP): Branch & Bound, Outer Approximation
– NP hard w.r.t. N



New approach: combine functional analysis arguments (convexification of velocity space, bang bang principle) with numerical approximation schemes – orders of magnitudes faster

Sager '06, Sager, B., Reinelt '08

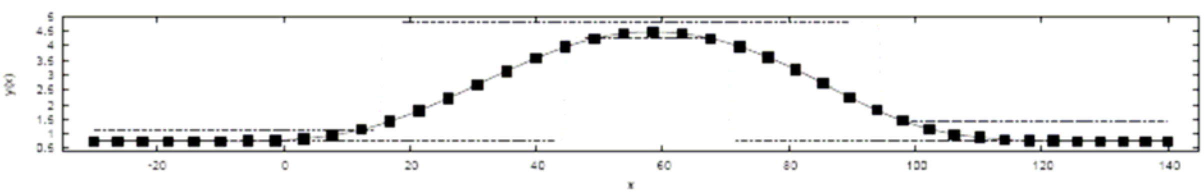
Example: Automobile Test Drive Problem

Automobile Test Drive Problem (Gerdts '05, '06)

Problem: ODE model for planar motion of an automobile

- 7 state variables, 4 controls
  - brake, accelerator position, steering wheel (continuous)
  - gear shift {1,2,3,4,5} (discrete)

- test problem 1: time optimal control, state constrained parcour



- results!

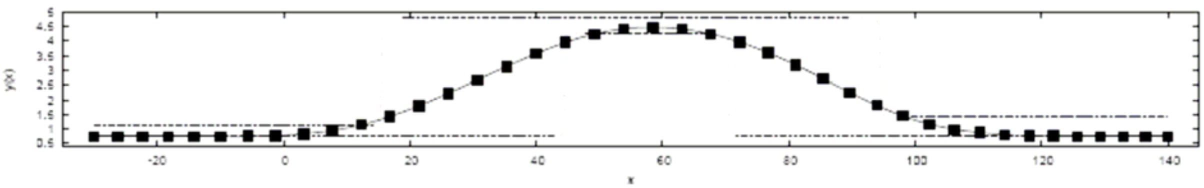
N	t_f	CPU Time
20	6.779751	00:23:52
40	6.786781	232:25:31
80	–	–

Automobile Test Drive Problem (Gerdts '05, '06)

Problem: ODE model for planar motion of an automobile

- 7 state variables, 4 controls
  - brake, accelerator position, steering wheel (continuous)
  - gear shift {1,2,3,4,5} (discrete)

- test problem 1: time optimal control, state constrained parcour



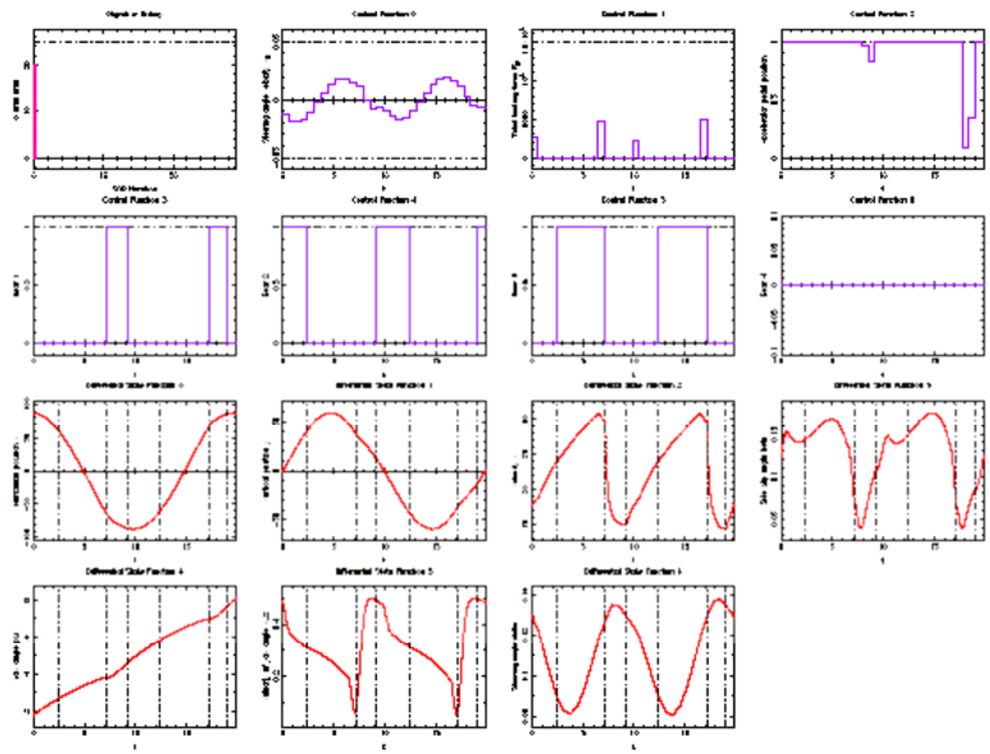
- results!

N	t_f	CPU Time	t_f	CPU Time
20	6.779751	00:23:52	6.779035	00:00:24
40	6.786781	232:25:31	6.786730	00:00:46
80	–	–	6.789513	00:04:19

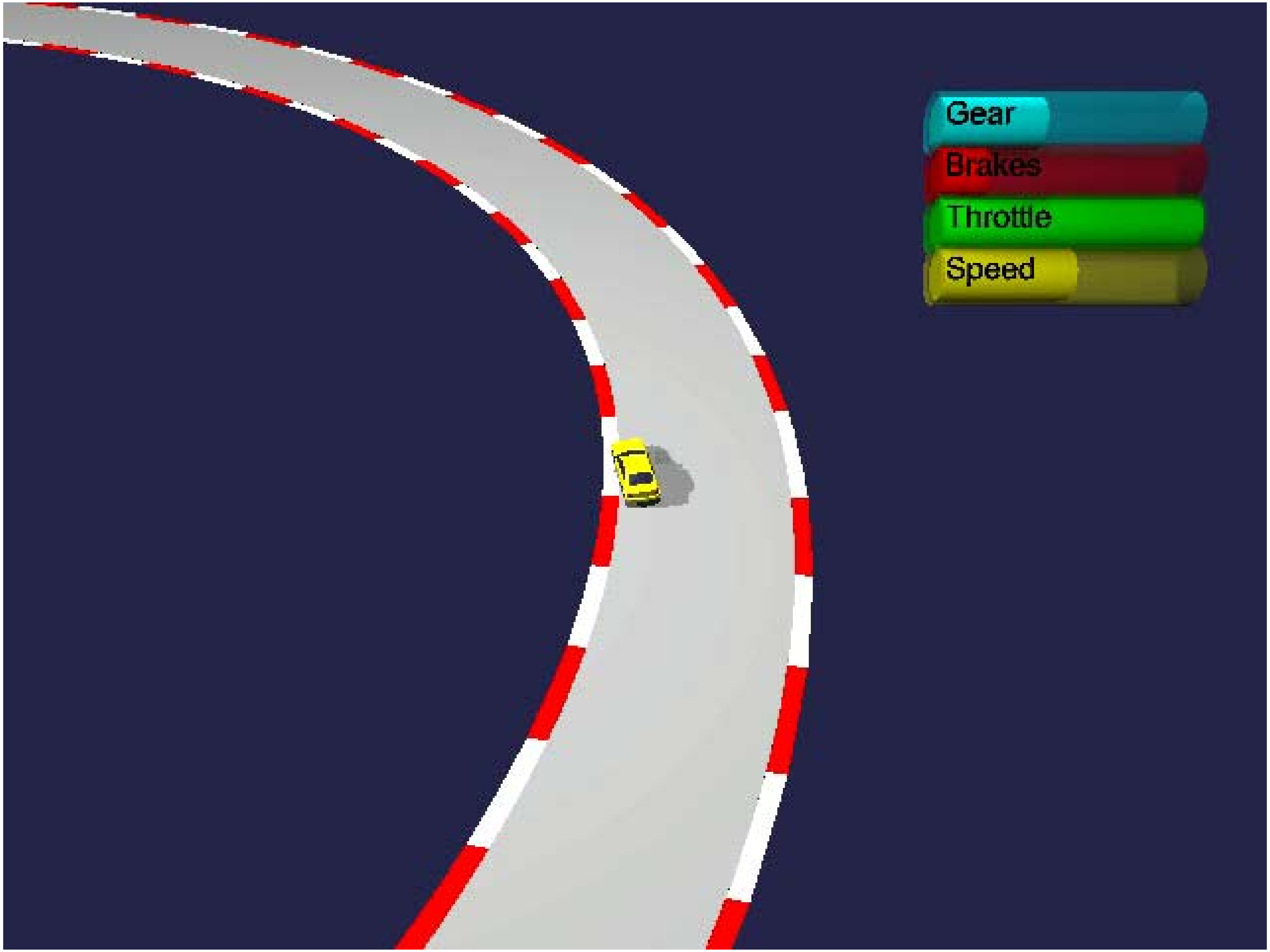
Kirches, Sager, B., Schlöder '08

Automobile Test Drive Problem

- test problem 2: time optimal control, state constrained ellipsoidal race track, periodic solution (Kirches, Sager '08)



Automobile Test Drive Problem (Kirches '08)



Challenges and Open Problems

- existence theory, estimates of # of switching points of controls and state equations
- sufficient optimality conditions including state constraints
- sensitivity and bifurcation analysis of parametric problems
- efficient algorithms for general class of hybrid problems
- computation of feed-back controls, generalizations to PDE

Requirements of Real-Time Optimization

NMPC: Nonlinear Model Predictive Control

Mathematical problem:

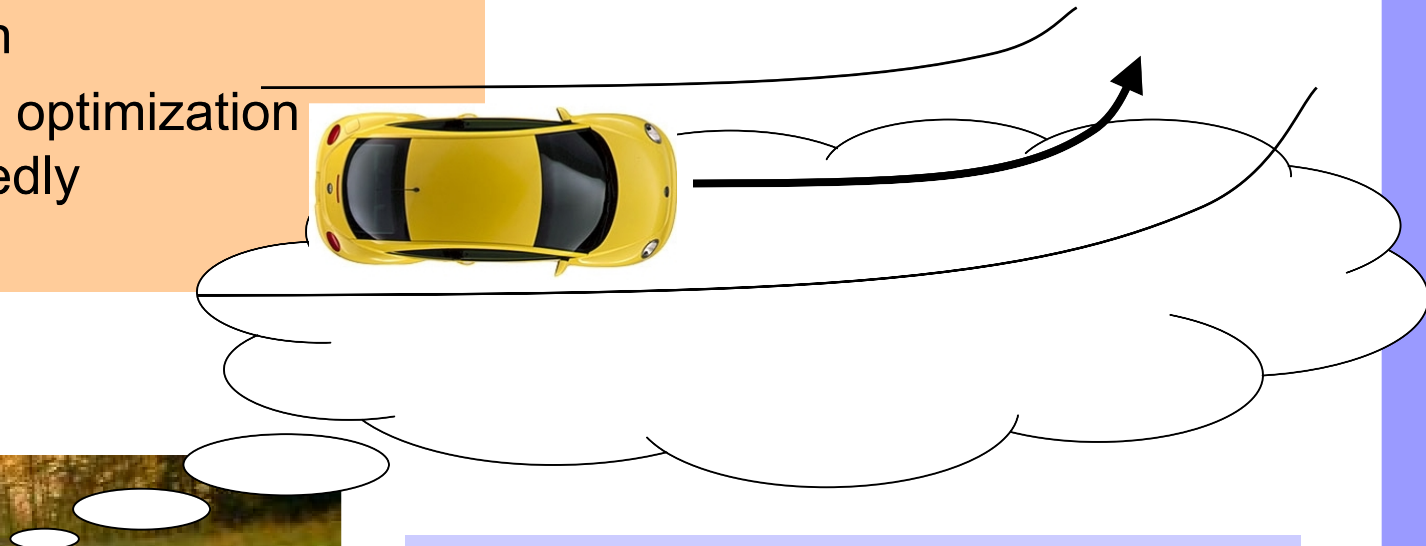
- real process deviates from model behaviour,
  - needs optimal feedback control $u(x^*(t_0))$
  - that responds as fast as possible to perturbations $x^*(t_0)$ of states

- "look ahead strategy": from the given perturbed state $x^*(t_0)$
  - solve a predictive optimal control problem
  - for a *"moving prediction horizon T"*
  - based on the nonlinear dynamical model
- choose first instant of optimal control as feedback control

NMPC: Nonlinear Model Predictive Control

look ahead strategy:

- detailed model allows prediction and optimization
- but must solve optimization problem repeatedly
- fast enough



Mathematical challenges:

- minimize control response time to perturbations
- need to solve many neighboring problems very fast

Mathematical Progress:
from Minutes to Microseconds

Classical approach:

- whenever new state estimates are available
    - start a fast Newton-type method like SQP that successively solves approximating quadratic programs
    - iterate to convergence (?), and then apply control solution to real process

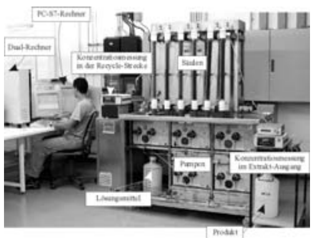
- New *nested* "real-time iteration" approach:
    - iterate *continually* – dovetailing the progress of dynamical process and optimization iterations
    - redefine optimization method such that in each iteration gradients, Hessians, quadratic programs can be computed as far as possible without knowing new states
    - insert actual state value at end of iteration and apply control response immediately

NMPC Applications

Distillation column (with Univ. Stuttgart)

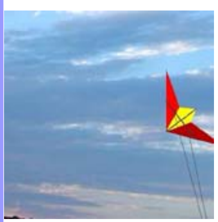
Polymeri-sation reactor (with BASF)

Combined Cycle Power Plant (with Univ. Pavia)



Chromatographic Separation (with Univ. Dortmund)

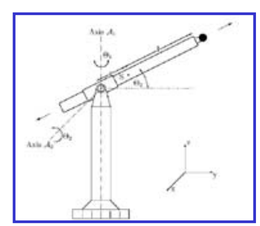
PET plant: Plant wide control project (with Politecnico di Milano)



Looping kites (with politecnico di Torino)

PREDIMOT, car engine control (with Univ. Linz, Stuttgart, Politecnico di Milano)

Oscillating chains test problem



Robot arms (with Columbia Univ.)

Example: Wind Power Generation Kites

Wind Power Generation Kites

Principle: use kites as alternative to wind mills

- power is generated when kite pulls out
- power is invested when kite is pulled in

Model considers:

- kite: lift, drag, and attitude
- wind shear and air density
- cable: elasticity, sagging, and drag

- periodicity, state and control constraints

Diehl, Ilshöfer, Houska '06

Wind Power Generation Kites

maximum power periodic orbit:

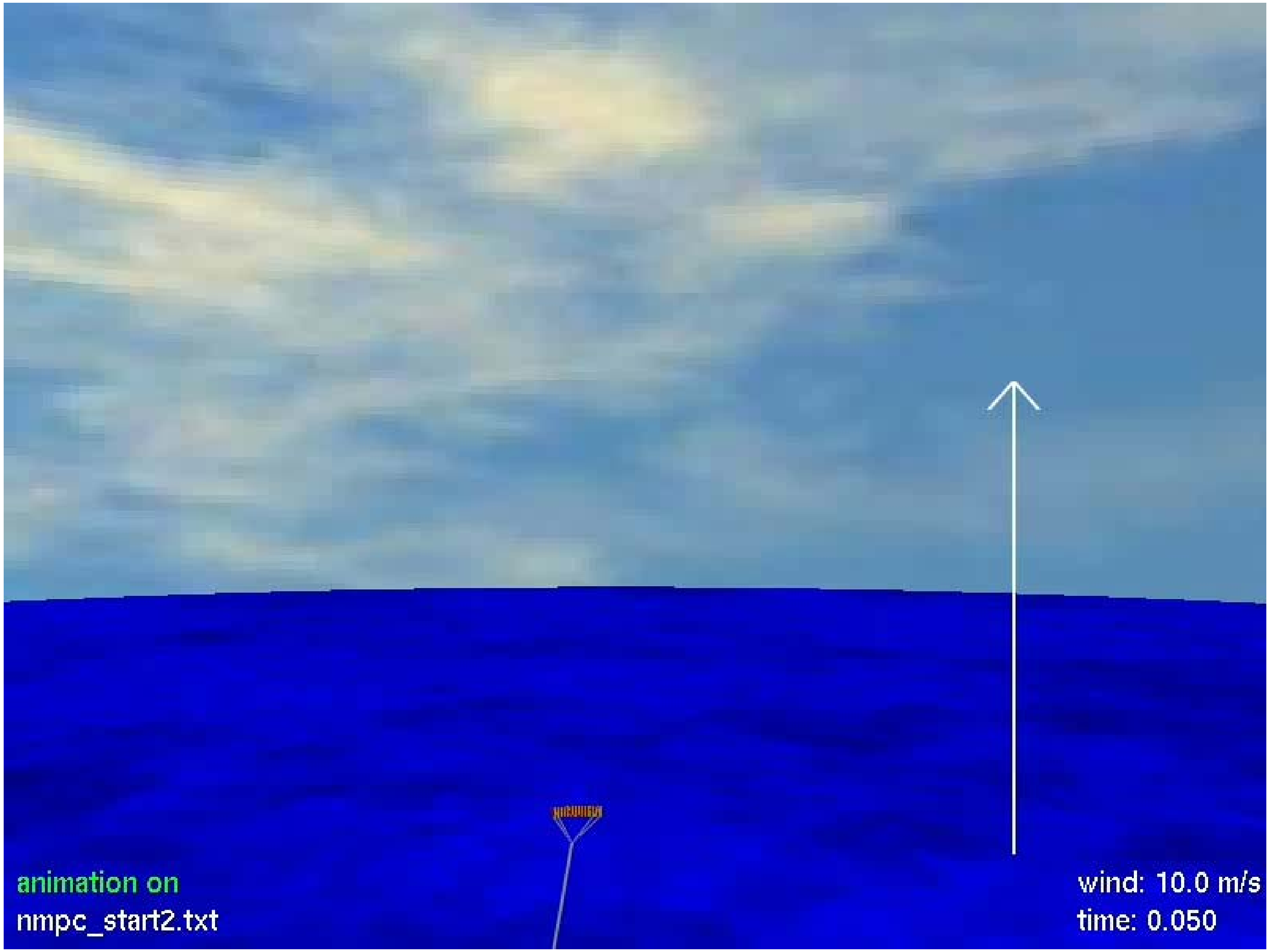
- for a kite with area of 500 square meters and wind speed of 10 m/sec:
  - optimal cable length: 1200 to 1300 meters
  - optimal cable diameter 6.7 cm
  - optimal cycle time 19.9 seconds

optimal average power production is 4.898 MW!

(for two kite system: up to 14 .13 MW!)

Wind Power Generation Kites

Optimal start-up trajectory into optimal periodic orbit without perturbations

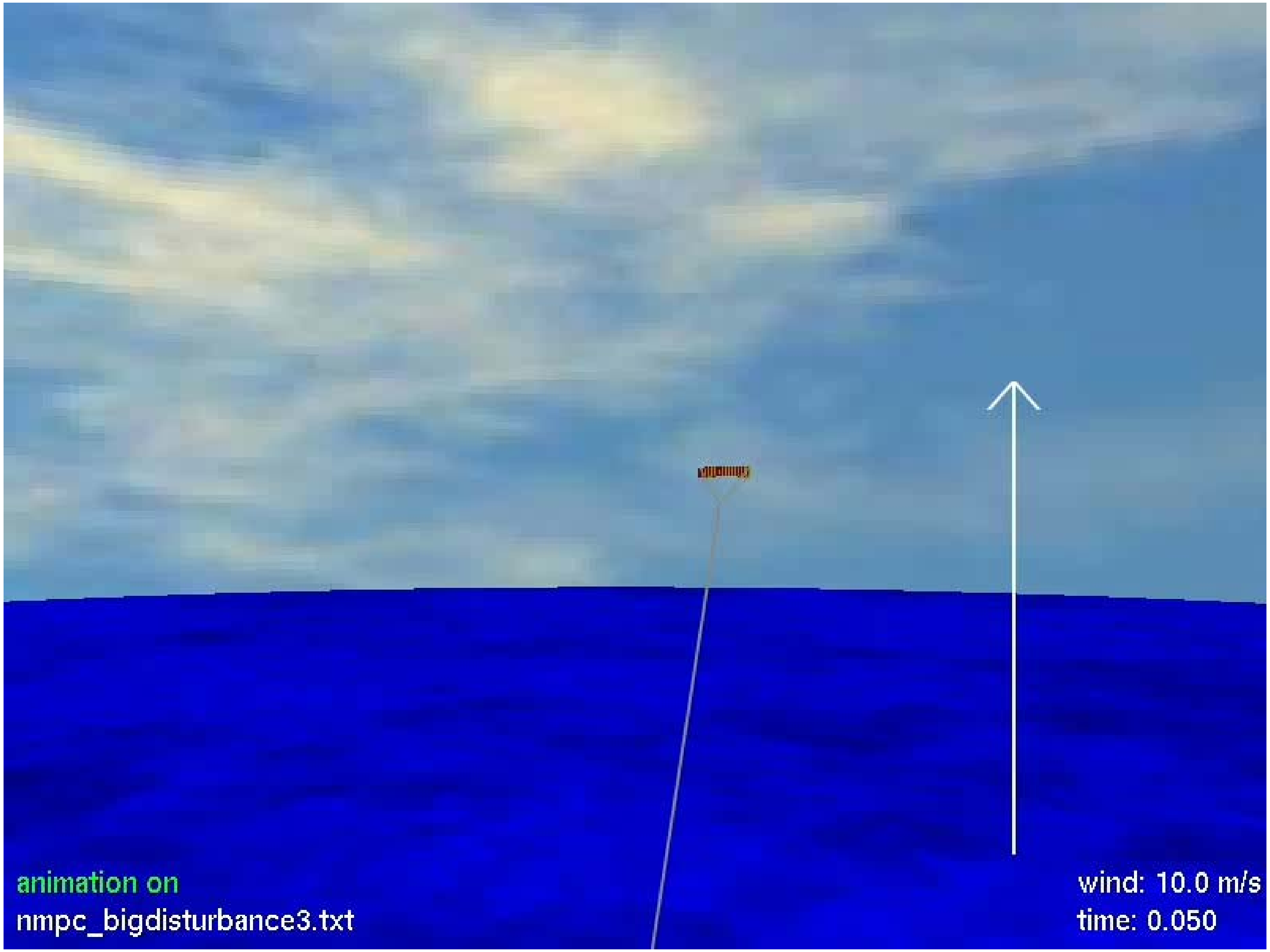


Wind Power Generation Kites: NMPC

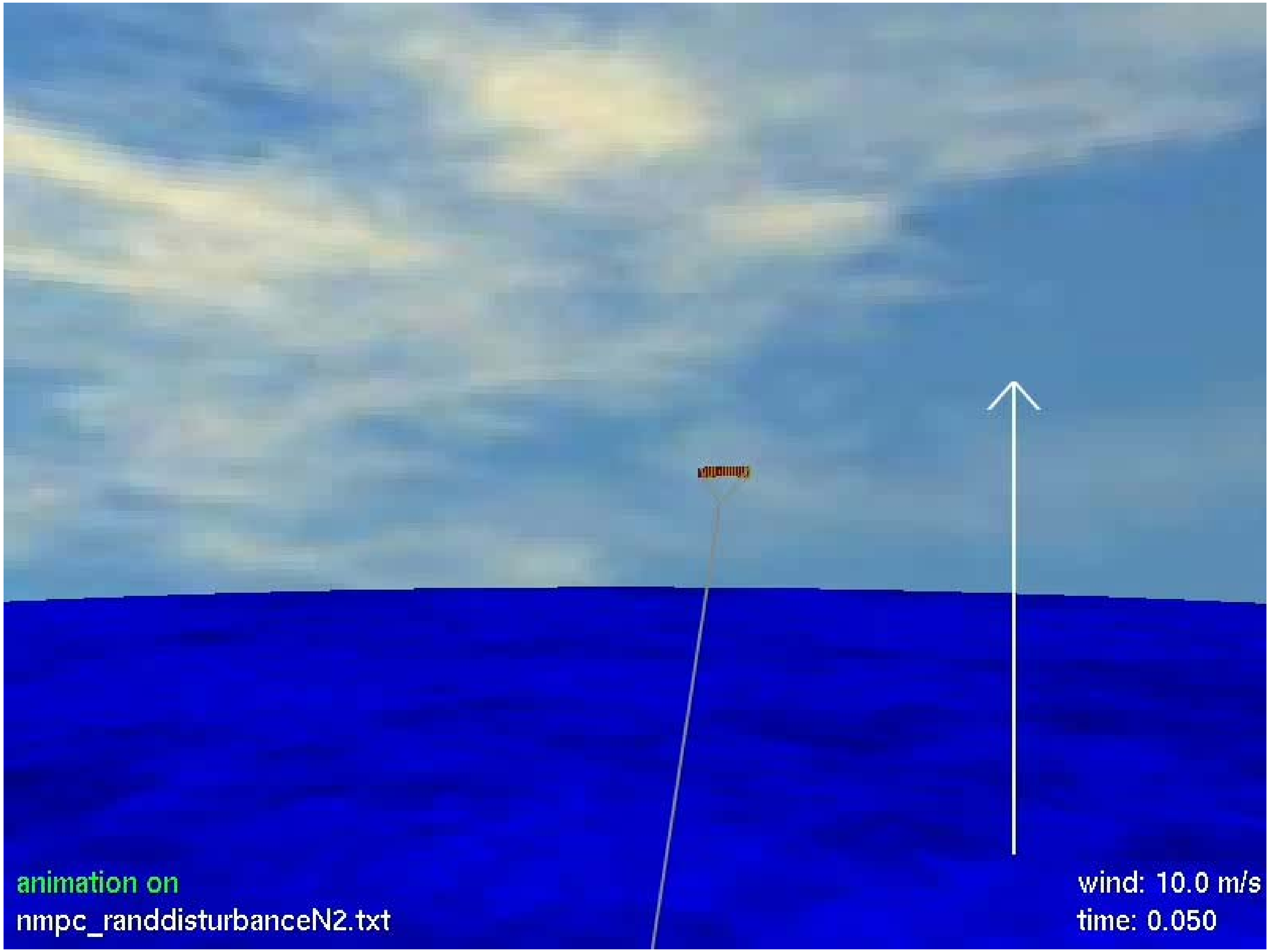
NMPC:

- tracking of periodic orbit under perturbations
  - mild: random wind +/- 2m/sec
  - strong: side wind gust 20 m/sec for 5 sec
- prediction horizon 10 sec
- sampling time: 1 sec
- real time iteration response time below 1 msec

Wind Power Generation Kites: NMPC



Wind Power Generation Kites: NMPC



Challenges and Open Problems

- "convergence" is unsuitable term under real time constraints
- so are "feasibility" and "optimality"
- proper analysis must rely on realistic measures of contractivity and asymptotic stability
- algorithm control must rely on error analysis and adaptivity
- almost completely lacking: analysis of different time scales of process and solution algorithms

Optimum Experimental Design for Dynamical Processes

Constrained Parameter Estimation Problem

Mathematical problem:

• determine parameters *p* and state variables *x* to minimize deviation from measurements η for suitable norm (e.g. weighted l_2, l_1, hybrid)

$$\min_{x;p} \Phi[x,p] = \left\| \eta - m[x,p] \right\|$$

• such that model equations and additional “experimental constraints” are satisfied

Constrained Parameter Estimation Problem

DAE process model

$$
\begin{aligned}
B(.)\dot{y} &= f(t,y,z,p,q,u) \\
0 &= g(t,y,z,p,q,u) \\
x &:= (y,z)
\end{aligned}
$$

data

$$
\begin{aligned}
\eta_{ij} &= b_{ij}(t_i, x(t_i), p, q) + \varepsilon_{ij} \\
\varepsilon_{ij} &\in N(0, \sigma_{ij}^2)
\end{aligned}
$$

determine p and x

$$
\min_{x,p} \sum_{i,j} w_{ij} \frac{(\eta_{ij} - b_{ij}(t_i, x(t_i), p, q))^2}{\sigma_{ij}^2}
$$

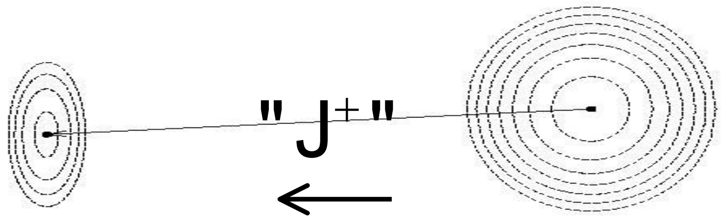
$$
\begin{aligned}
B(.)\dot{y} &= f(t,y,z,p,q,u) \\
0 &= g(t,y,z,p,q,u) \\
d(x(t_0),\ldots, x(t_f), p, q) &= 0, \quad \text{or } \geq 0
\end{aligned}
$$

Assessment of Statistical Error of Estimate

- good fit is not enough!

- must assess uncertainty of state and parameter estimate $X^*(e)$ - prediction quality! - depending on measurements errors, e.g.

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



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

- accuracy of parameter and state estimate given up to first order by a generalized inverse J^+, resp., covariance matrix C

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

Example: Enzyme Reaction Kinetics

Enzyme Reaction Kinetics

- enzymes = industrial biocatalysts
- expensive evaluation of long-term stability behavior
- in practice: many expensive experiments are carried out

$$
\begin{tikzcd}
N \arrow[rr, shift left, "k_{NU}"] \arrow[dr, "k_{ND}"'] & & U \arrow[ll, shift left, "k_{UN}"] \arrow[dl, "k_{UD}"] \\
& D &
\end{tikzcd}
$$

N: native enzyme

U: unfolded enzyme

D: deactivated enzyme

Enzyme Reaction Kinetics

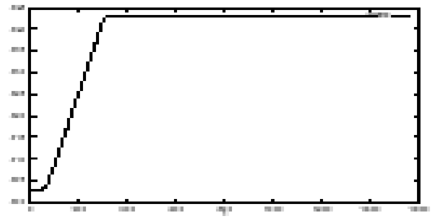
degussa.
creating essentials

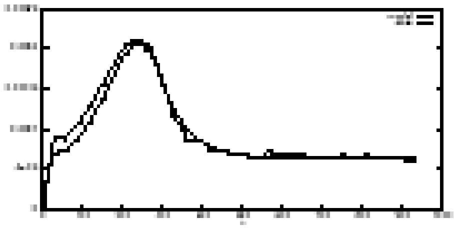
- 4 state variables, (non-linear) Arrhenius kinetics
- 8 unknown reaction rate parameters p
- 1 control function u(t): temperature
- 1 indirect measurement device only: amount of base added to neutralize acid production (side-reaction)
- wanted: stability quantities "total turn-over", "half-life"
- will see: problem very ill-conditioned, cannot estimate parameters from a single "standard" experiment

Experiments with *Candida Antarctica* on Ionic Resin ("Novozym")

degussa.
creating essentials

- estimated values of parameters $\pm$ standard deviations after one standard experiment
- fit looks okay, but ...





	initial profile
p_1	27.86 ± 4.42
p_2	48.98 ± 10.92
p_3	$1.73 \pm 2.39 \times 10^5$
p_4	$634.20 \pm 806.00 \times 10^6$
p_5	$-1.43 \pm 1.50 \times 10^7$
p_6	$-7.50 \pm 4.16 \times 10^7$
p_7	-4.15 ± 0.091
p_8	-8.63 ± 2.00

- ... parameter estimates mostly useless

Can We Optimize Experimental Conditions to Improve Accuracy?

aim:

- choose optimal experimental conditions $\xi = (u, q, w)$
  - u: control functions: temperature profiles, feed-streams,
  - q: control parameters: volume, initial conditions,
  - w (**integer!**): sampling design: measurement devices and times

cost function

- **maximize information gain, here: minimize uncertainty of resulting parameter estimate**

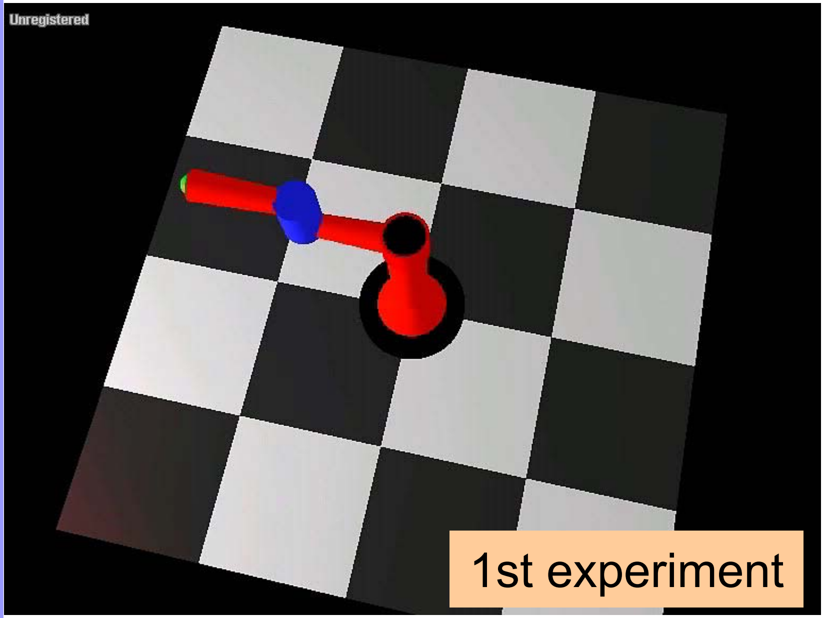
- subject to state, control and parameter constraints, e.g.
  - costs, safety, feasibility (shifts), domain of model validity (!),...

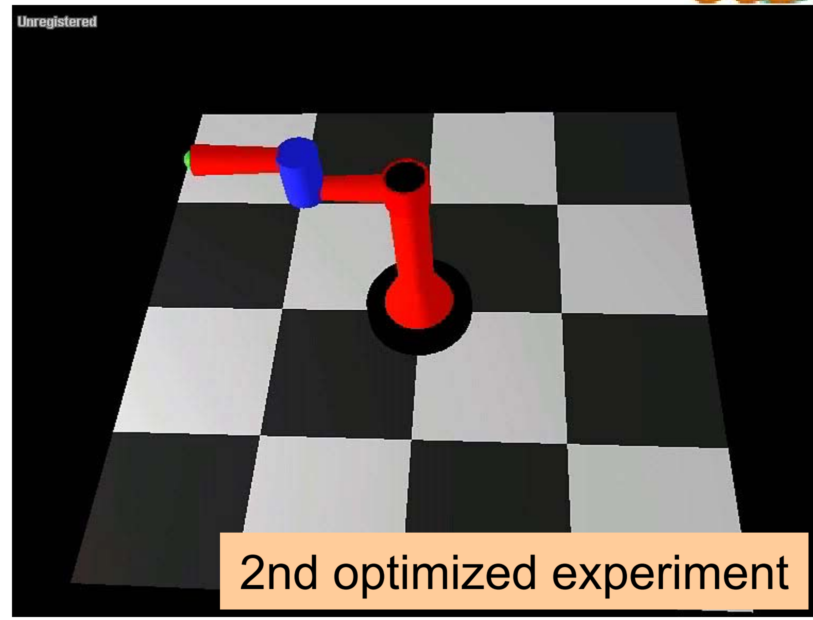
Example: Calibration of Robots

Example: Calibration of Robots

- off-line programming of robots requires precisely calibrated models
- dynamic parameters (e.g. inertias) and kinematic parameters only approximately known
- high repetition accuracy, but different for individual robots
- simplest case: identify 2 inertial parameters for SCARA robot from position measurements of tool center point

Example: Calibration of Robots





Result: one additional optimal experiment improves parameter accuracy by one order of magnitude - equivalent to 99 repetitions of original experiment!

very high optimization potential indeed!

Optimum Experimental Design is a Complex Non-Standard Optimal Control Problem

- improve accuracy of estimate by minimizing a function of the covariance matrix

$$\Phi(C(X,\xi))$$

$$\text{e.g. } \Phi(C) := \text{trace } C \text{ (A - criterion), or max } C_{ii} \text{ - properly weighted}$$

e.g. Kiefer, Wolfowitz, Box, Draper, Pukelsheim, ...

- subject to model equations, state and control constraints

$$c(X,\xi) \geq 0$$

- cost function
  - already *involves 1st order derivatives* of model
  - *implicity* defined via a generalized inverse
  - *non-separable* - no Pontryagin's maximum principle

Optimum Experimental Design for Discrimination between Model Candidates

e.g. Fedorov, Atkinson

- design experiments to reject models by lack-of-fit test
- leads to complex multiple experiments, multiple models optimal control problems
- to maximize "difference" between predicted measurements of model candidates
- but fortunately "standard" cost functions, and highly structured

e.g. Dieses '97, Hoffmann '06

Numerical Methods for Optimum Experimental Design

- direct approach: discretization of control problem
- RSQP solution of constrained NLP, exploitation of structures
- efficient evaluation of 1st and 2nd derivatives

- sequential approach: generate sequence of optimized experiments, exploit information of the previous

VPLAN/OPTEX

Bauer, Körkel, B., Schlöder '98, '02

- robust "worst case" optimum experimental design: make experiments insensitive to parameter uncertainties

Körkel, B., Kostina, Schlöder '04

Experiments with *Candida Antarctica* on Ionic Resin ("Novozym")

degussa.
creating essentials

- 1, 2, 3 or 4 ***simultaneously* optimized experiments** do not allow to identify all reaction rate coefficients up to an order of magnitude
- this application requires 5 simultaneously optimized experiments!

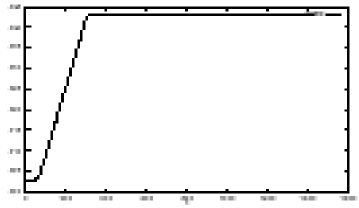
	initial profile	5 designed + 1 initial profiles
p_1	27.86 ± 4.42	27.69 ± 0.81
p_2	48.98 ± 10.92	49.96 ± 1.97
p_3	$1.73 \pm 2.39 \times 10^5$	0.548 ± 0.075
p_4	$634.20 \pm 806.00 \times 10^6$	184.88 ± 26.10
p_5	$-1.43 \pm 1.50 \times 10^7$	-4.32 ± 0.26
p_6	$-7.50 \pm 4.16 \times 10^7$	-6.20 ± 2.20
p_7	-4.15 ± 0.091	-8.88 ± 1.59
p_8	-8.63 ± 2.00	-11.34 ± 7.54

estimated values of rate coefficients and standard deviations after parameter estimation from experimental data

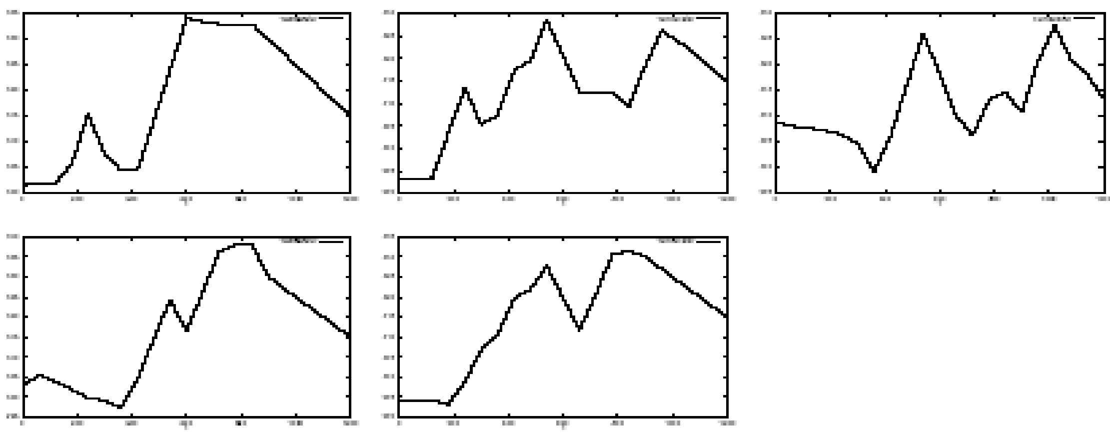
Experiments with *Candida Antarctica* on Ionic Resin ("Novozym")

- the 5 "simultaneously optimal" temperature control profiles:

  - initial ("standard") temperature control profile

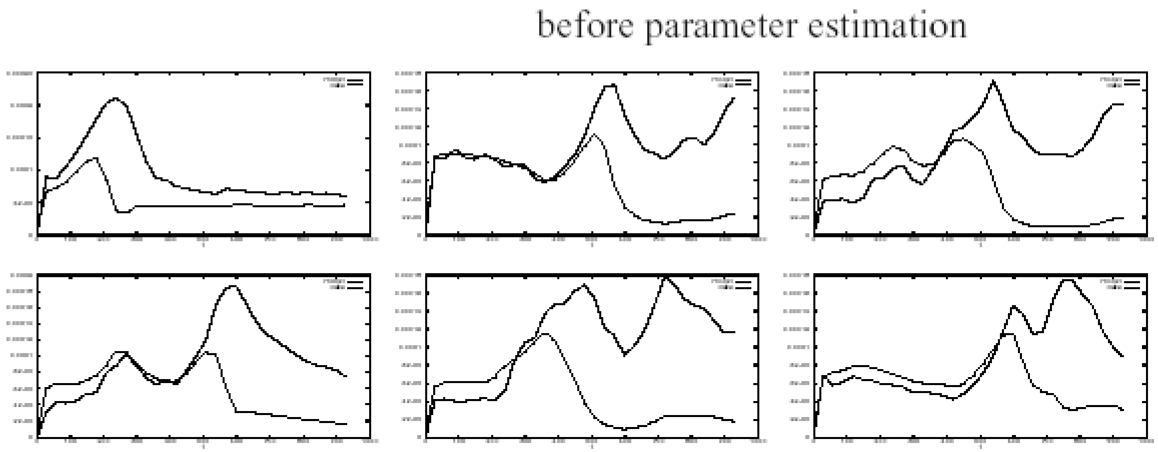


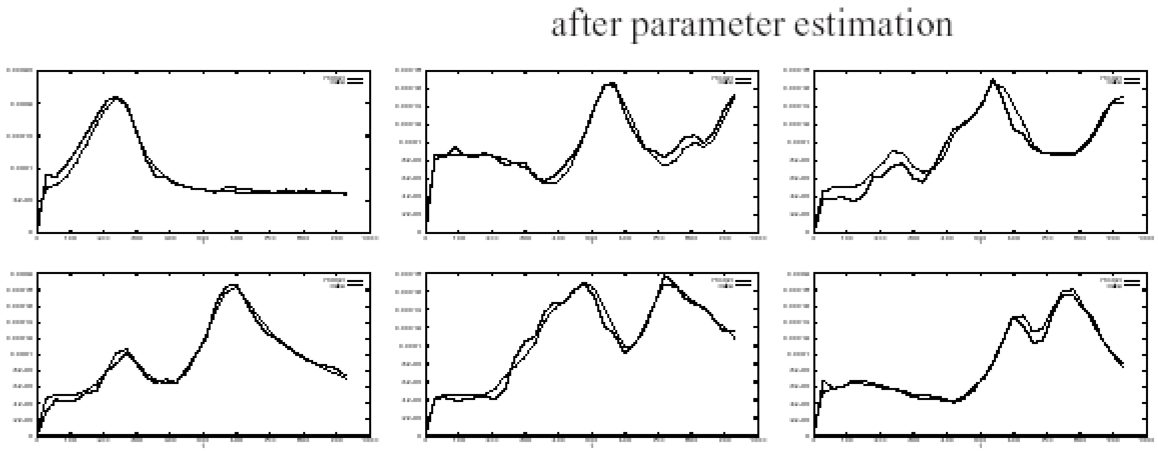
  - plus 5 additional "simultaneously optimal" profiles



Experiments with *Candida Antarctica* on Ionic Resin ("Novozym")

- computational results also deliver good fit:





***Robust* Optimum Experimental Design**

Aim: Reduce Sensitivity w.r.t. Parameter Uncertainty

Robust Optimum Experimental Design

problem: parameters p only known up to - large - confidence region!

- desirable: worst case design - expensive!

$$\min_{\xi \in \Omega} \max_{\underbrace{\|\Sigma (p - p_0}_{\text{confidence}} \underbrace{)\| \leq \gamma}_{\text{region!}}} \Phi(C(\xi, p))$$

- approach chosen: 1st order approximation

$$\min_{\xi \in \Omega} \max_{\|\Sigma(p-p_0)\|_2 \leq \gamma} \left(\Phi(C(\xi, p_0)) + \nabla_p \Phi(C(\xi, p_0))^T (p - p_0) \right)$$

- yields "robust" cost function and constraints (γ, Σ "natural" weights)

$$\min_{\xi \in \Omega} \left(\Phi(C(\xi, p_0)) + \gamma \left\| \Sigma^{-1} \nabla_p \Phi(C(\xi, p_0)) \right\|_2 \right)$$

- generalizes to general optimization problems under uncertainty

Körkel, B., Kostina, Schlöder '00, '04

Diehl, B., Kostina '06

Enzyme Kinetics: Robust vs. Nonrobust Design

- simplified reaction scheme
- sequential design for 4 reaction rate coefficients
- pseudo measurement data generated by numerical experiments

"true"
27.77
50.15
0.55
185.25

first estimate	
26.0455	$\pm$ 0.371394
45.6499	$\pm$ 0.973549
5.58527	$\pm$ 4.4236
1885.01	$\pm$ 1495.38

the first "expert", but non-optimized experiment

Enzyme Kinetics: Robust vs. Nonrobust Designs

- parameter estimation for one initial experiment followed by
- sequence of *one* additional optimal experiment after another

9 nonrobust   3 robust

27.4901 ± 0.187754	27.825 ± 0.27924	27.77
49.4083 ± 0.48714	50.2741 ± 0.724924	50.15
0.569797 ± 0.0249352	0.543646 ± 0.0385069	0.55
191.798 ± 8.35222	182.532 ± 12.8385	185.25

conclusion: enormous potential - drastic reduction of experimental costs

Challenges and Open Problems

- rich and well advanced theory of optimum experimental design for linear models (from Kiefer and Wolfowitz to Draper, Fedorov, and Pukelsheim)

- hugh potential, but practically no theory for nonlinear dynamical processes available

- also more efficient numerical methods needed
  - more general problem classes, e.g., hybrid systems
  - on-line optimum experimental design problems
  - incorporation of OED in real time state and parameter estimation

- **very high relevance in industrial applications, BASF donates 200 k€/year for a Junior Research Group in OED at the new Graduate School of IWR**

Summary

- detailed models provide the engineers with a powerful tool to optimize and control complex dynamical processes
- potential can only be realized by powerful new mathematical methods,
  - neither by trial and error,
  - nor by sophisticated heuristics,
  - nor by textbook mathematics
- there is a wealth of engineering problems in industry and technology waiting for new analytical and computational methods

and maybe the cow

will learn to fly!