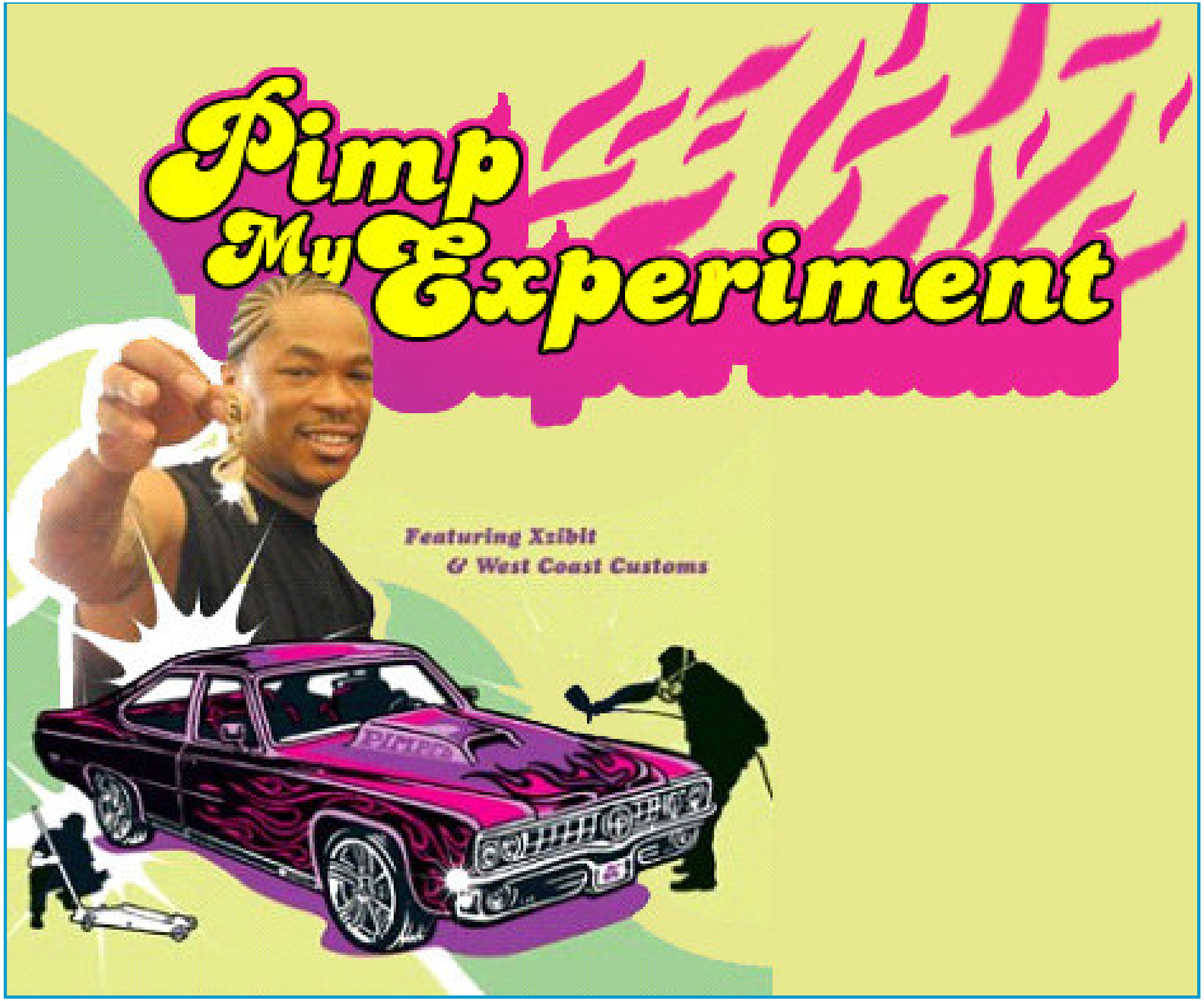
The METER approach for the optimal experimental design of ill-posed problems … and more

André Bardow
Department of Process & Energy
Delft University of Technology

OPTEX: Workshop on Optimal Experimental Design in Engineering
October 8-9, 2007
Katholieke Universiteit Leuven

Outline

- Optimal design of diffusion experiments
  an application case study
- METER designs for ill-posed problems
  a novel problem formulation


Pimp
My Experiment
Featuring Xzibit
& West Coast Customs

Pimp my diffusion experiment

take an old car…

make it faster, nicer,… – better !

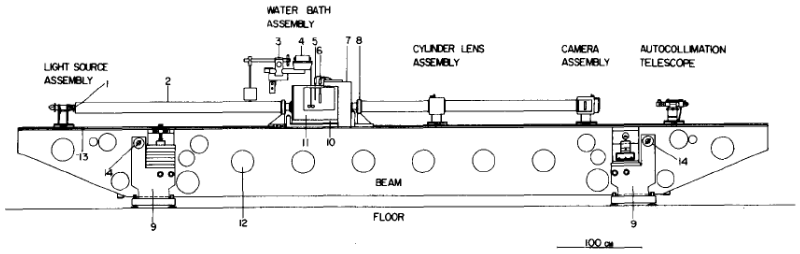
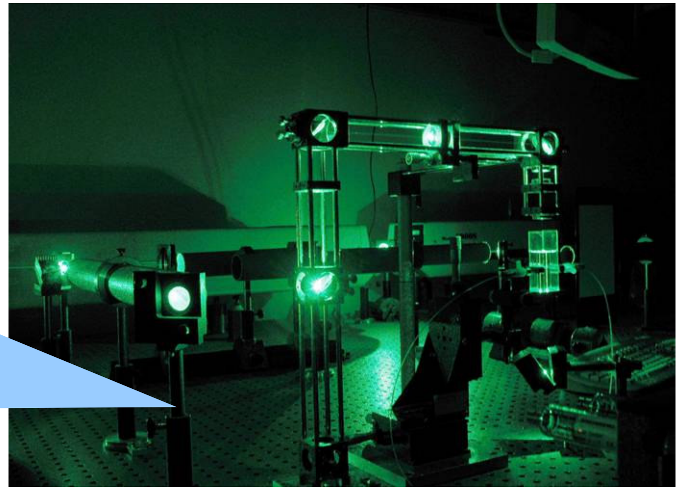


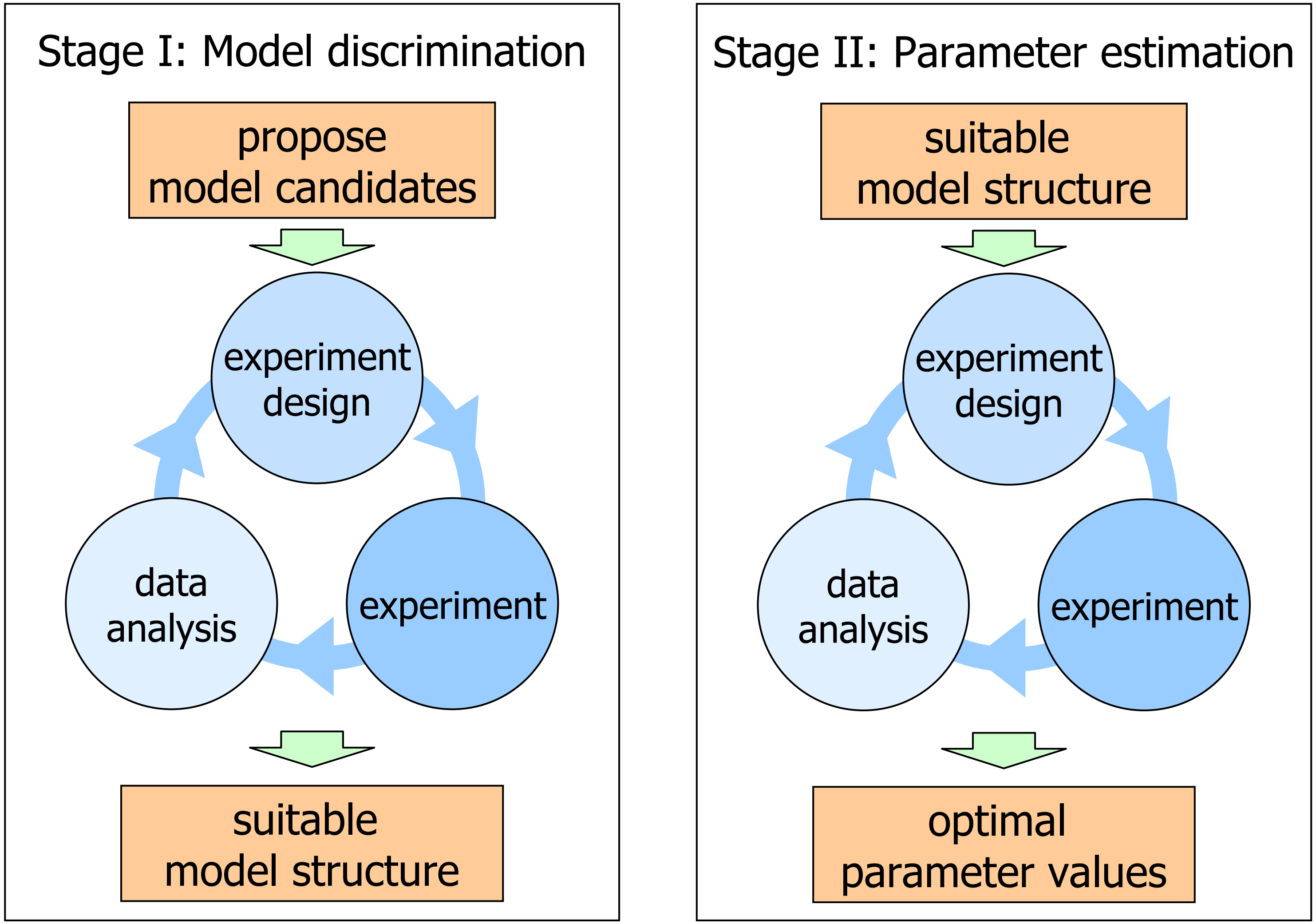
FIG. 1. Schematic drawing of the diffusiometer. 1—Slit of light source; 2—tube to minimize air currents in the light path; 3—stirring motor mount; 4—stirring motor; 5—sensor of temperature controller; 6—heater; 7—cell frame support; 8—special window rider; 9—beam support; 10—water bath support; 11—water bath; 12—access hole to reach the inside of the beam; 13—ways; 14—leveling collar to check any change in the position of beam relative to that of beam support.

take an old experiment…



make it faster, nicer,… – better !

Model-based experimentation



OED for parameter estimation

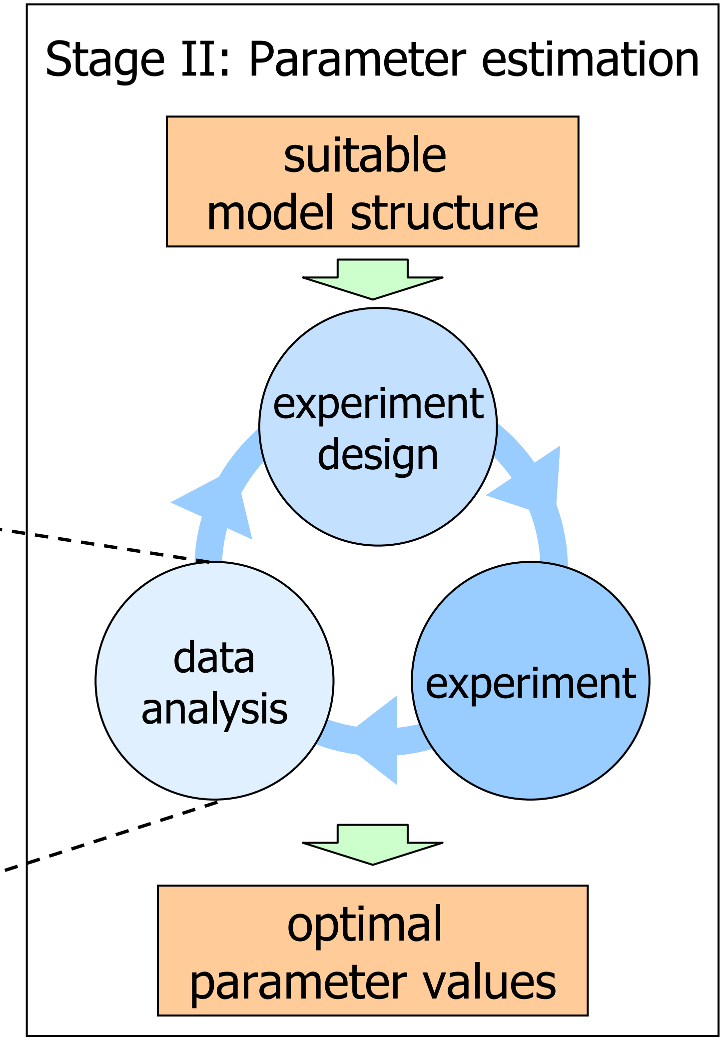
- Starting point: $y = f(x, \theta)$
  known model structure but unknown parameters
- Goal: determine parameters with high accuracy

Parameter estimation problem

model - measured

$$\min_{\theta} \sum_{\mu=1}^{N} \frac{\left[y(x_\mu, \theta) - \tilde{y}(x_\mu) \right]^2}{\sigma^2(x_\mu)}$$

Least-squares-type objective



Optimal Experimental Design: Idea

set free variables ***x***
such that information
on parameter θ is maximized

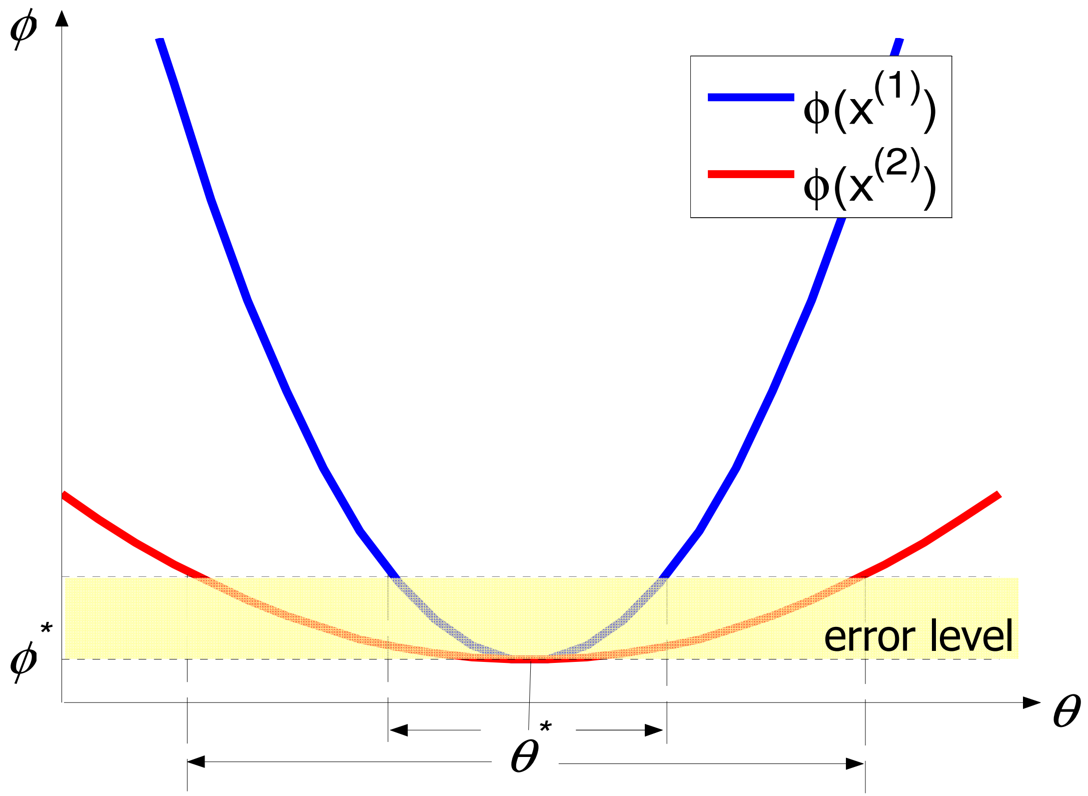
=

maximize curvature of
parameter estimation objective

OED for parameter estimation:

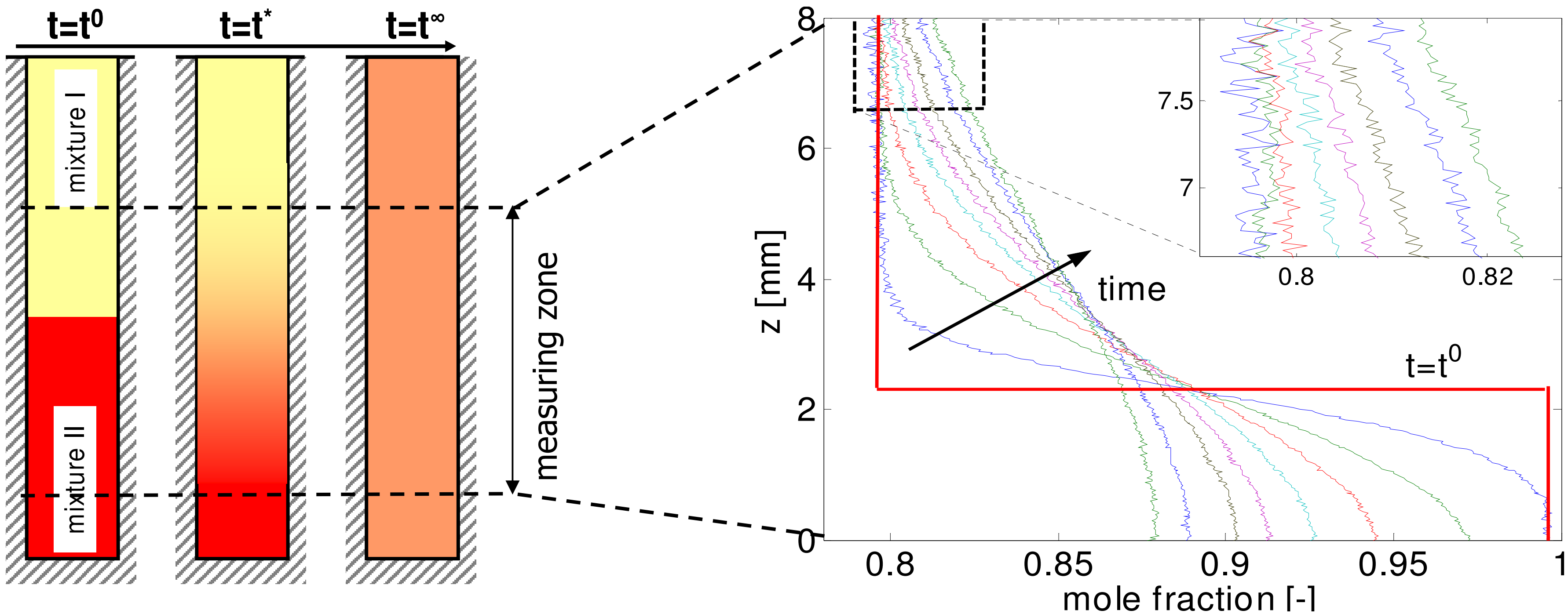
$$\max_x \frac{\partial^2 \phi}{\partial \theta^2}$$

parameter estimation objective



Case study: Raman diffusion experiments

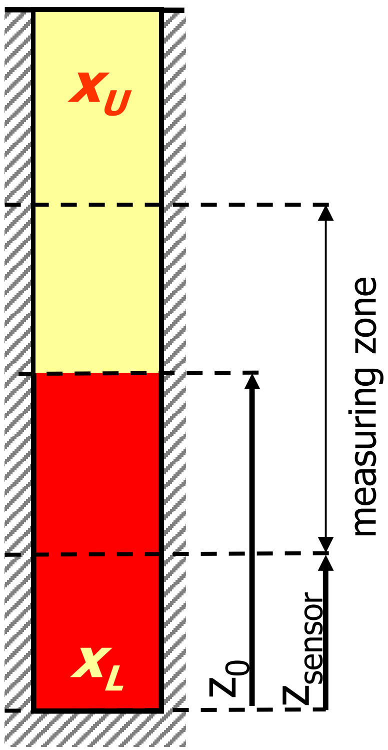
(experiments: Göke, Oertker, Koß, LTT, RWTH Aachen)



- analysis of one-dimensional diffusion
- nonlinear calibration approach (Alsmeyer et al., 2003,2004)
- simultaneous measurement of all mole fractions
- high resolution (Δt = 10 s, Δz = 20 μm)
- measurement error: statistical ≤ 0.2 mol-%, systematic ≤ 0.5 mol-%,

(Bardow et al., AIChE J., 2003, 2006)

Design questions



→ what mixture volume ratio?

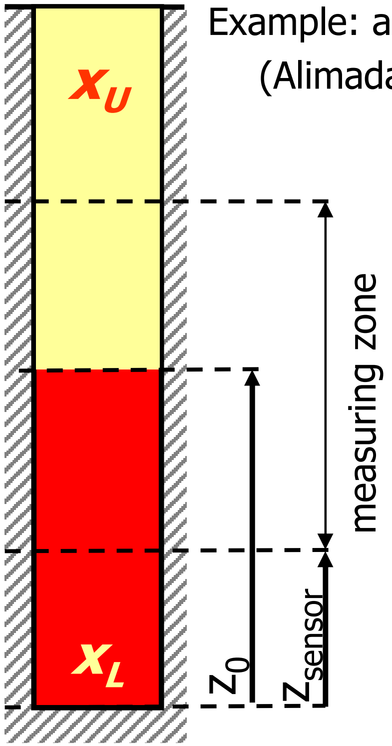
→ where to measure?

→ how long to measure?

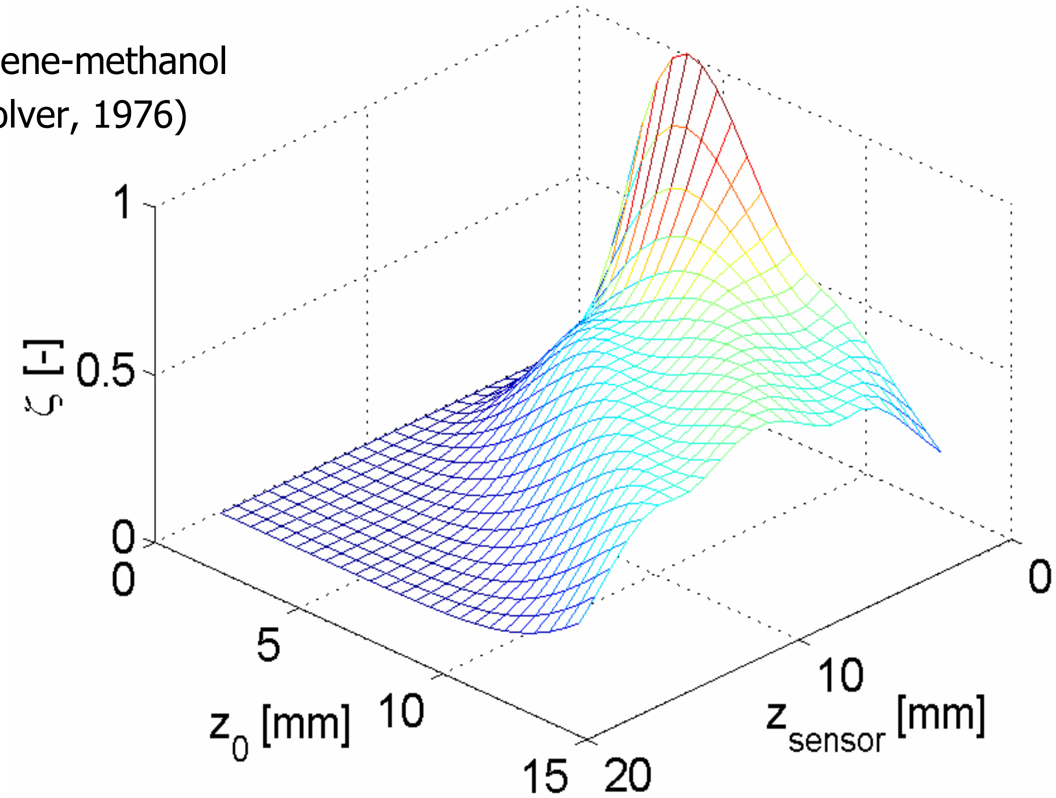
→ which mixture compositions?

optimal experimental design for diffusion coefficient measurements

Optimal Experimental Conditions

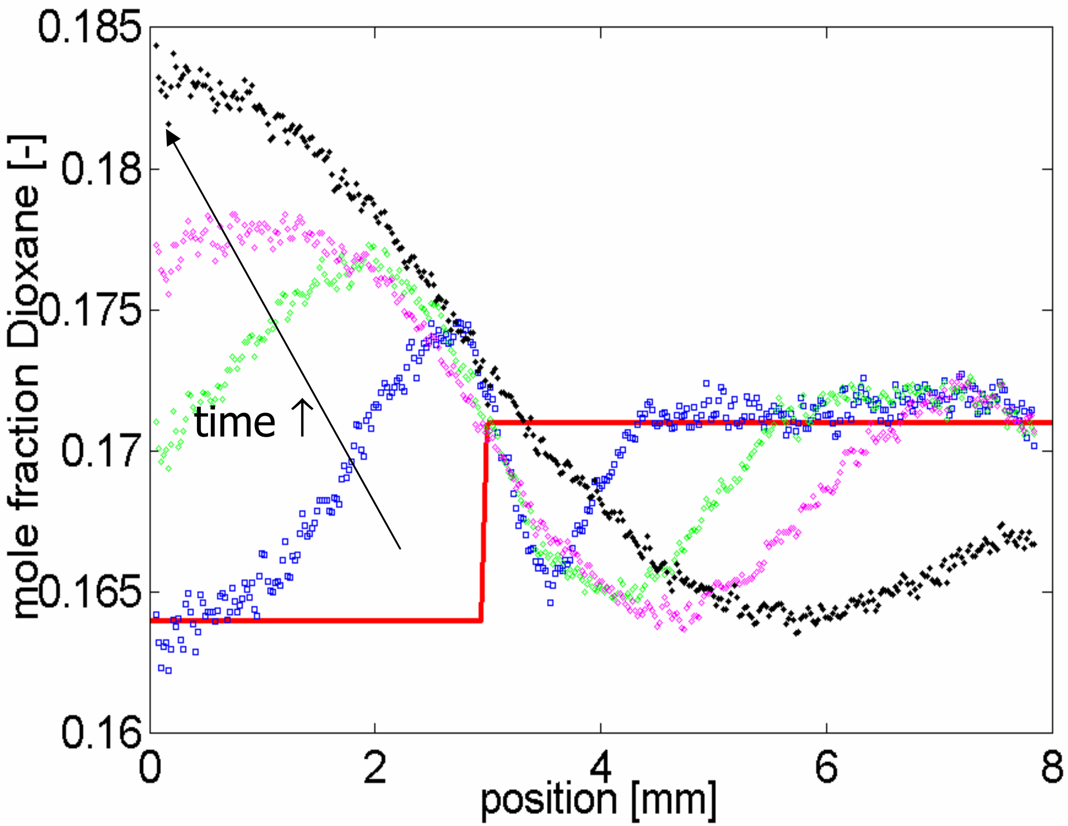


Example: acetone-benzene-methanol
(Alimadadian and Colver, 1976)



- scaled objective ζ-efficiency measures information per parameter
- measurements at the wall, *i.e.*, restricted diffusion experiments
- unequal volume of both phases, actual value depending on molar volume
- optimal experiment duration depends on cell size → optimal Fourier number

Wall effect in ternary system



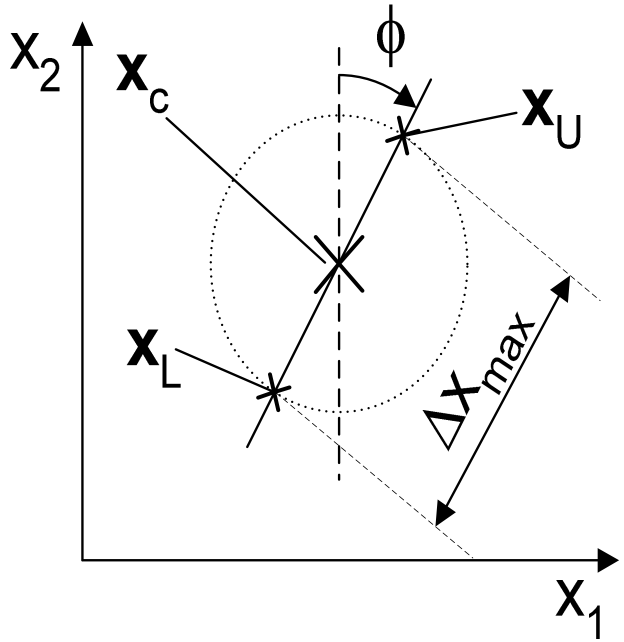
Toluene-Dioxane-Cyclohexane;

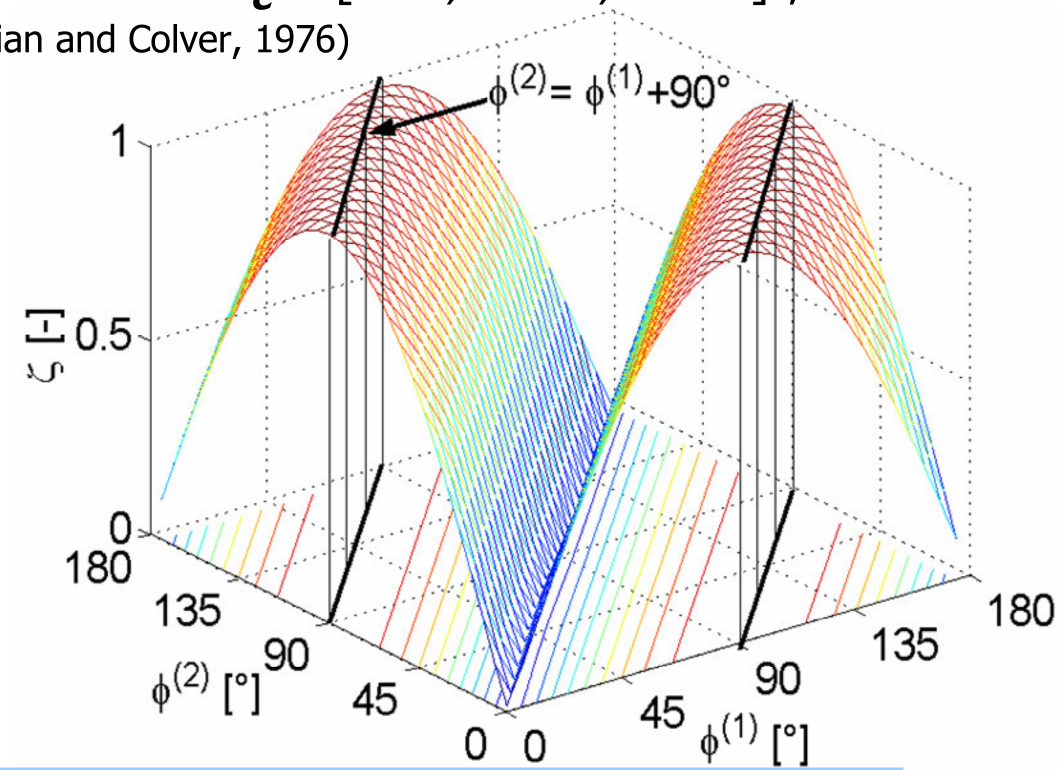
$x_{Dioxane} \approx$ const

confirmation of optimal design prediction in Raman experiments

Optimal Ternary Mixture Composition

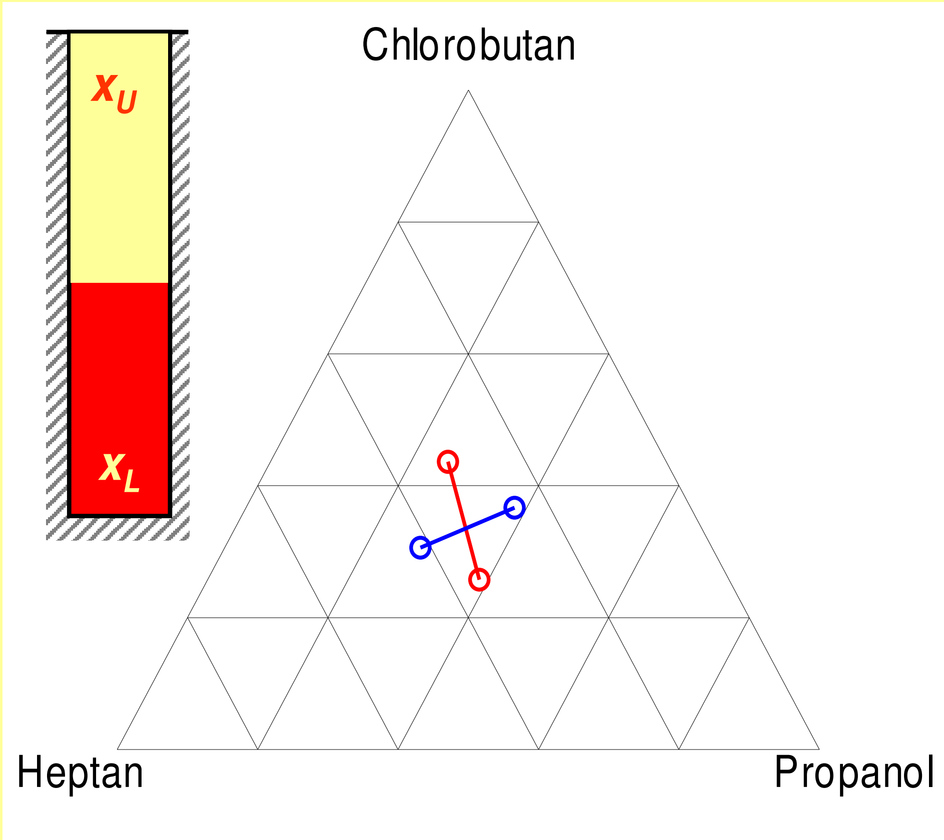
Example: acetone-benzene-methanol at $\mathbf{x}_c = [0.35; 0.302; 0.348]^T$,
(Alimadadian and Colver, 1976)





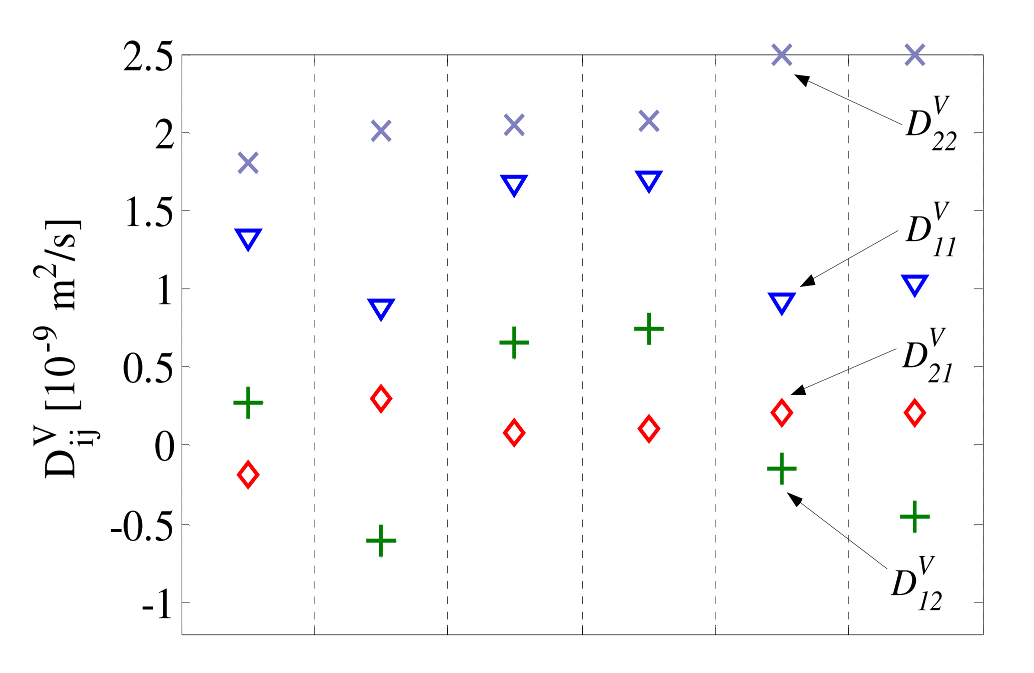
- one Raman experiment suffices to determine ternary Fick matrix
- two experiments leads to substantial improvement
- experiments should be as distinct as possible ($\phi^{(2)}=\phi^{(1)}+90°$)
- include constraints to ensure hydrodynamic stability

Ternary Raman experiment



Ternary Results

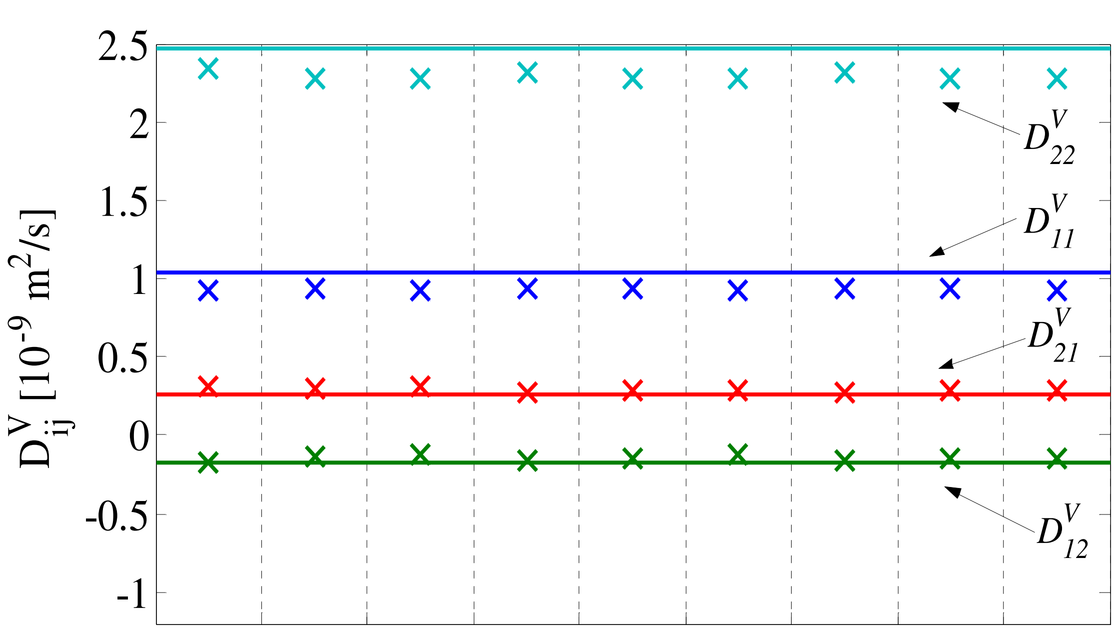
Diffusivities from a single Raman experiment



→ one Raman experiment gives full diffusion matrix
→ currently scatter in data is still significant

Ternary Results

Diffusivities from two ***optimal*** Raman experiments



with reference data from Weingärtner et al. (1994)

→ one Raman experiment gives full diffusion matrix
→ good precision from 2 optimized runs
→ robust & efficient measurement
→ quantitative validation of design predictions

Summary: Diffusion experimtents

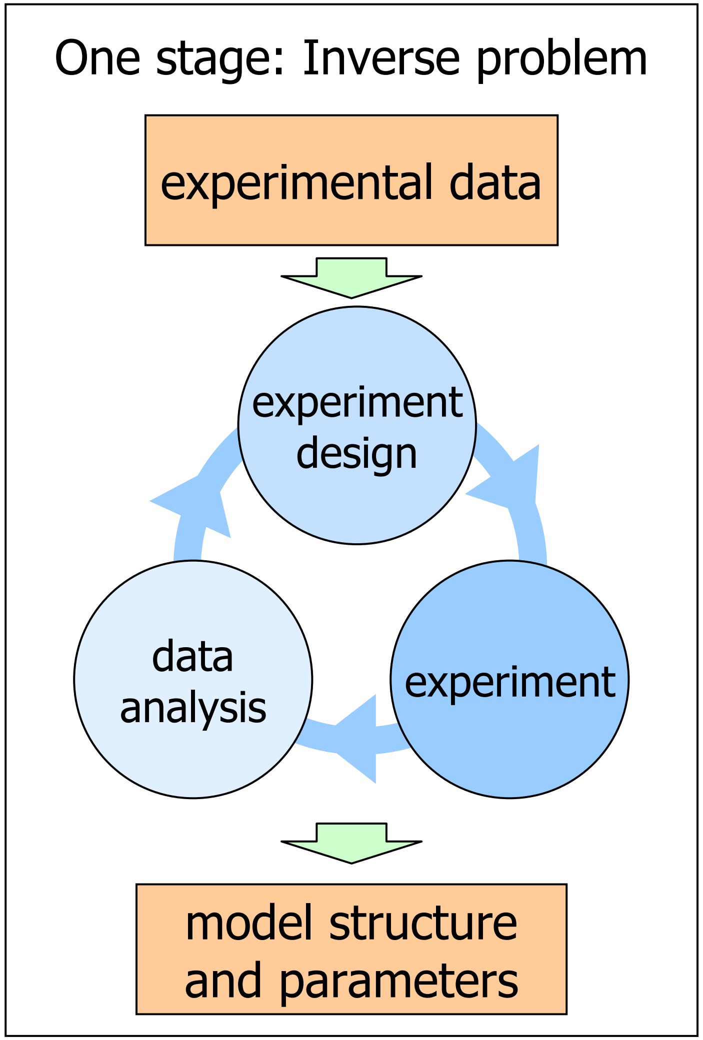
- OED **improved accuracy** by one order of magnitude
- OED led to **non-trivial problem insight**
- OED led to **generic design rules** for diffusion experiments

→ establishment of a **truly multicomponent diffusion** experiment

→ **efficient development** due to use of model-based approach

→ model-based analysis of **interferometry** and **Taylor dispersion**

Model-based experimentation



Two stage approach:
"Divide & conquer"

☺ for *a priori* knowledge

☹ restricted search space

One stage approach:

☺ explore full search space

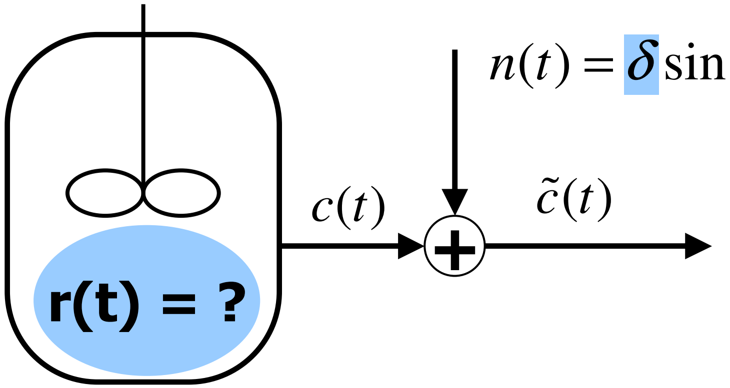
☺ no *a priori* knowledge required

→ higher requirements on data

→ tougher estimation problem

Ill-posed Posedness of Estimation Problem

Illustrating example: Batch reactor



a problem is well-posed if

- a solution exists
- the solution is unique
- small errors in data lead to small errors in the solution

Hadamard (1923)

Estimation problem

$$\min_r \left\| c - \tilde{c} \right\|$$

$$s.t.\ \ c(t) - c_0 = \int_0^t r(t)\,dt$$

but…

- arbitrary large errors for bounded δ

$$\left\| r(t) - \hat{r}(t) \right\|_\infty = \left\| \omega\delta \cos(\omega t) \right\|_\infty$$

- errors governed by error level and frequency
- estimation errors do not depend continuously on data

problem is ill-posed

TU Delft

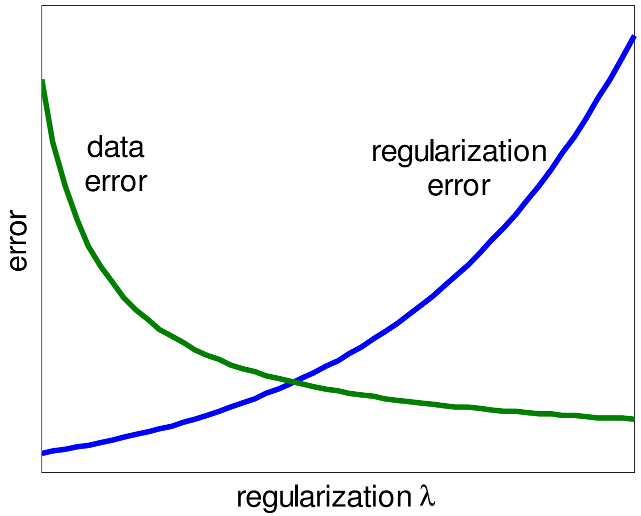
Tikhonov regularization

Estimation problem

$$\min_r \|c - \tilde{c}\|$$

$$s.t.\;\; c(t) - c_0 = \int_0^t r(t)dt$$

$$g(t) = \int_T K(t,s;d) f(s) ds$$



regularization

Regularized problem

$$\min_r \|c - \tilde{c}\| + \lambda \|r\|$$

$$s.t.\;\; c(t) - c_0 = \int_0^t r(t)dt$$

$$\min_f \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(g(t_i) - \int_T K(t_i,s;d) f(s) ds \right)^2 + \lambda \|Lf\|_{L_2}^2$$

lack of fit ("least-squares") → variance

smoothness penalty → bias

bias-variance trade-off crucial for success

Optimal Experimental Design: Idea

set free variables ***x*** such that parameter variance is minimal

||

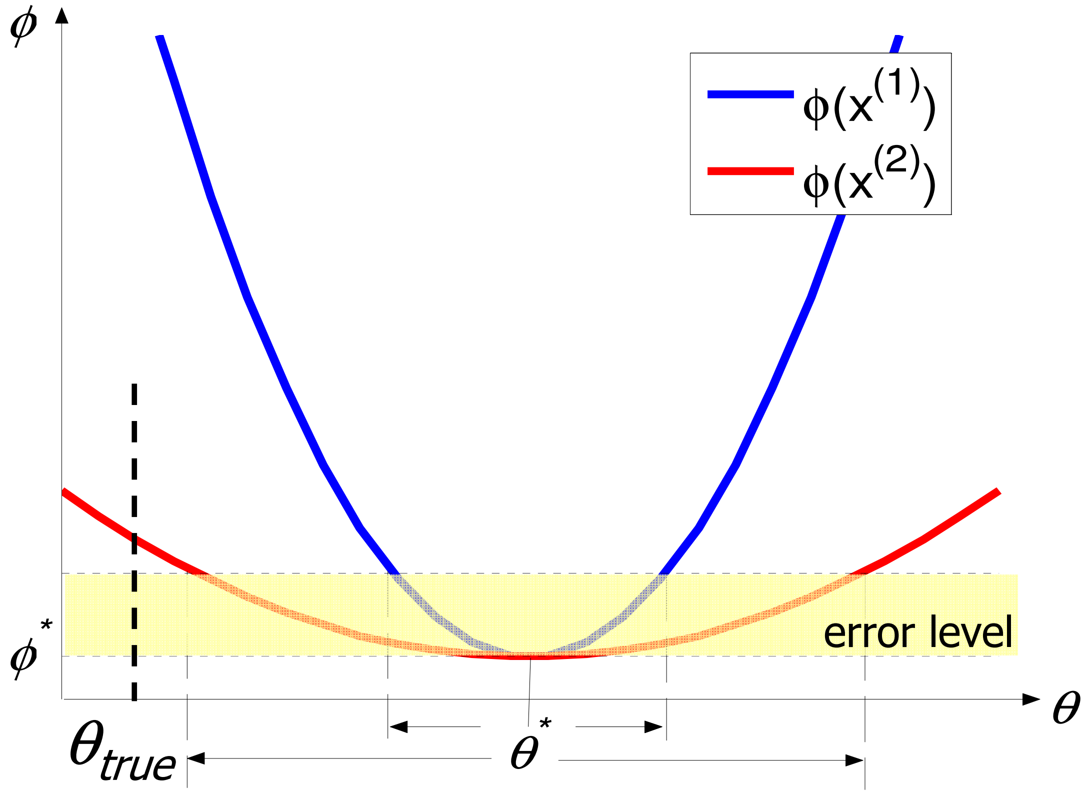
maximize curvature of parameter estimation objective

OED for parameter estimation:

$$\max_x \frac{\partial^2 \phi}{\partial \theta^2}$$

- focus on variance only
- bias contribution is missing
- not applicable

parameter estimation objective



"estimate the wrong solution with less uncertainty"

Minimum Expected Total Error (METER)

Optimal design:

find settings for experiment that allow best determination of unknown function

classical approach (Box-Lucas school): minimize parameter variance only

approach proposed for ill-posed problems (Box-Draper school):

minimize

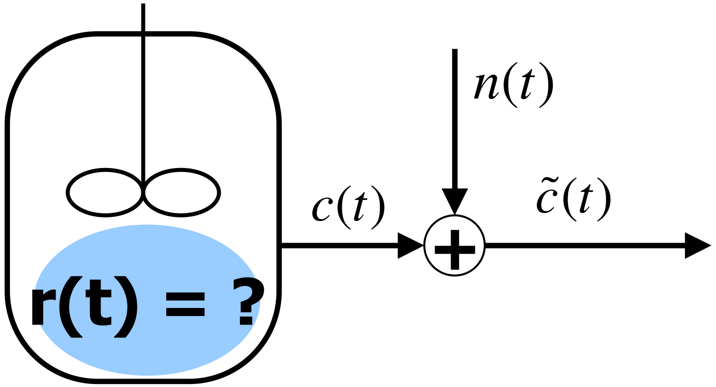
$$E\left(\left\| f - \hat{f} \right\|_{L_2}^2\right)$$ expected total error in estimate

$$= \left\| f - K^+(\lambda) K f \right\|_{L_2}^2$$ = *bias* due to regularization

$$+ \sigma^2 trace\left(K^+(\lambda)\left(K^+(\lambda)\right)^T\right),$$ + *variance* in estimate

with $K^+(\lambda) = \left(K^T K + \lambda L^T L\right)^{-1} K^T$. "regularized Fisher information matrix"

Batch Reactor Example



- determine reaction rate *r(t)*
  from concentration measurements *c(t)*

$$c(t) = \int_0^1 r(t')dt' \quad \Rightarrow \quad r(t) = \frac{dc(t)}{dt}$$

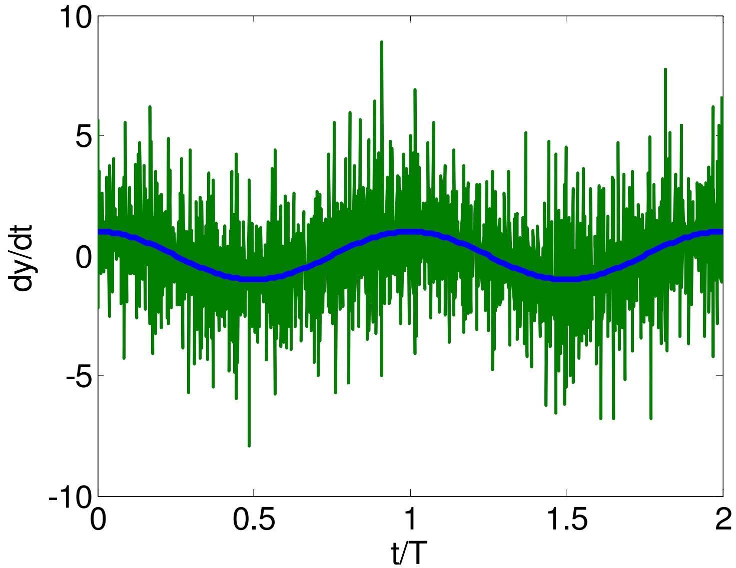
- numerical differentiation of experimental data

$\Rightarrow$ ill-posed problem

- solution by finite differences

$$\hat{r}(t_i) = \frac{c(t_i) - c(t_{i-1})}{\Delta t}$$

- time step Δt too large $\rightarrow$ approximation error
- time step Δt too small $\rightarrow$ amplification of noise

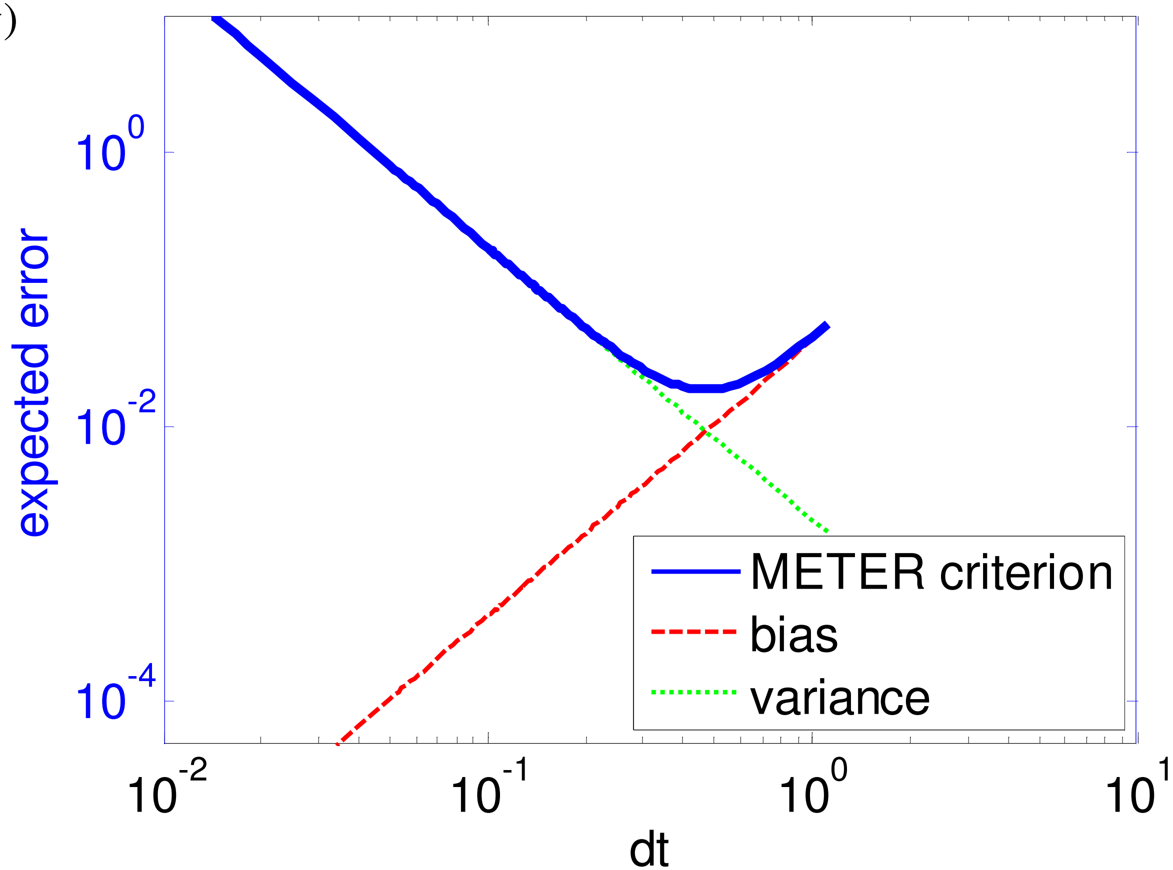


determine best time step in optimal design

Optimal sampling time by METER Criterion

$r(t) = \exp(-t)$

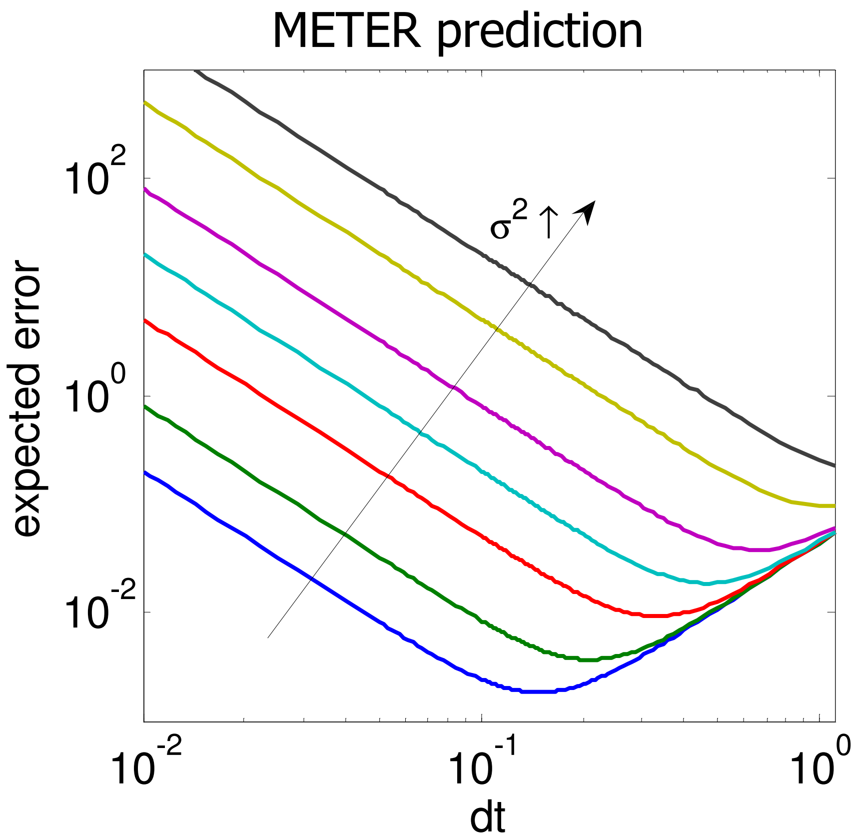
$t = [0,10]$

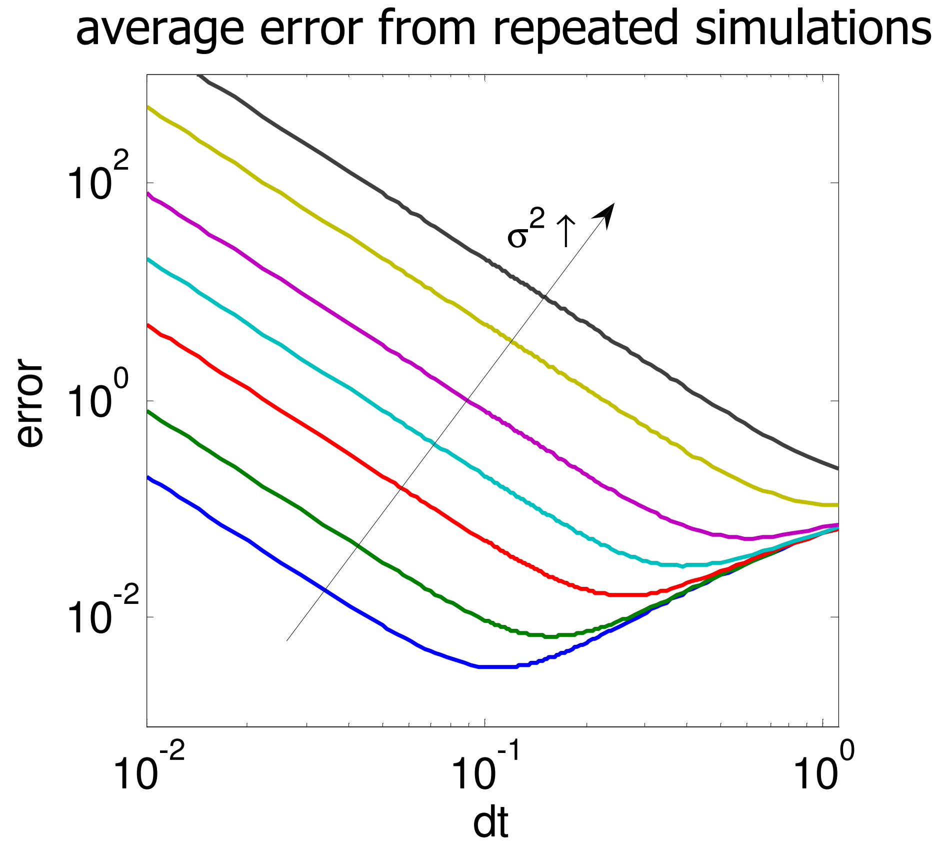


- bias and variance show the expected trends
- optimal sampling time as trade-off balance by noise level

TU Delft

Comparison with Numerical Computations





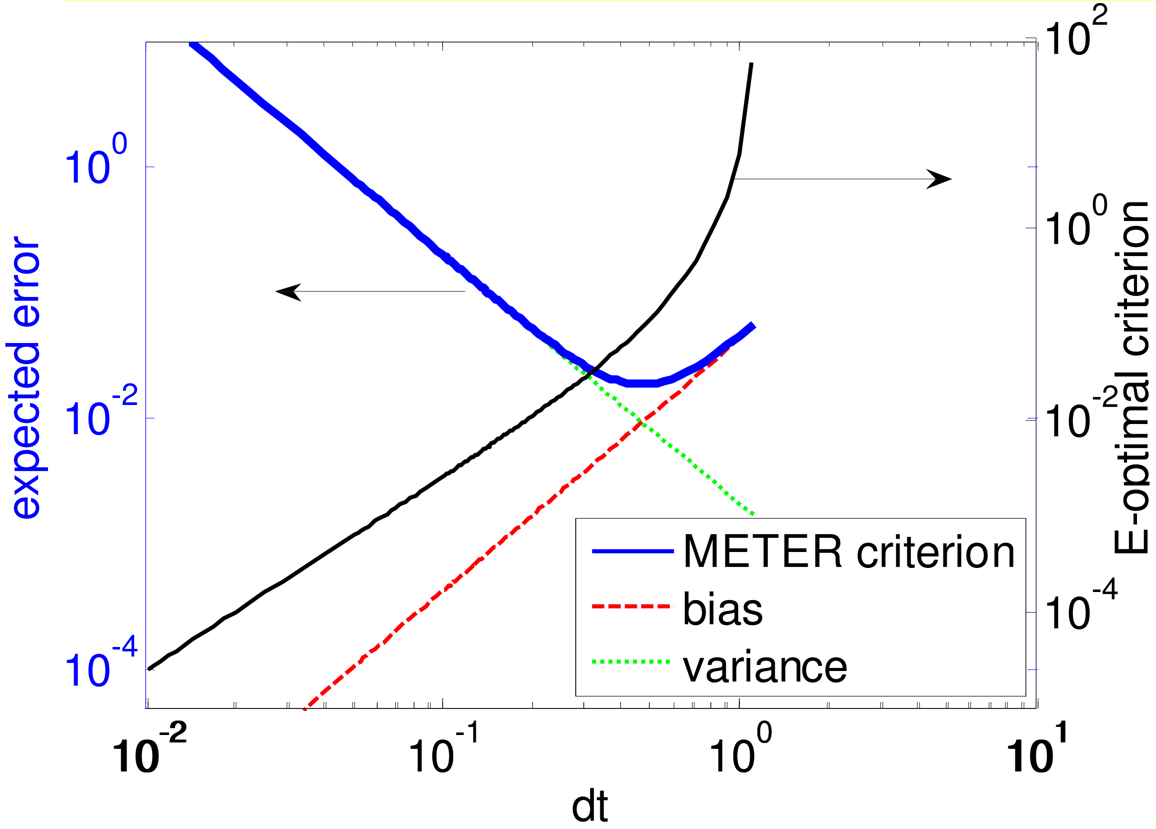
- predicted behavior is found in practice
- small difference from approximate bias term

Comparison with Classic Criteria

$r(t) = \exp(-t)$

$t = [0,10]$

E-optimal design: max min eigenvalue of Fisher matrix



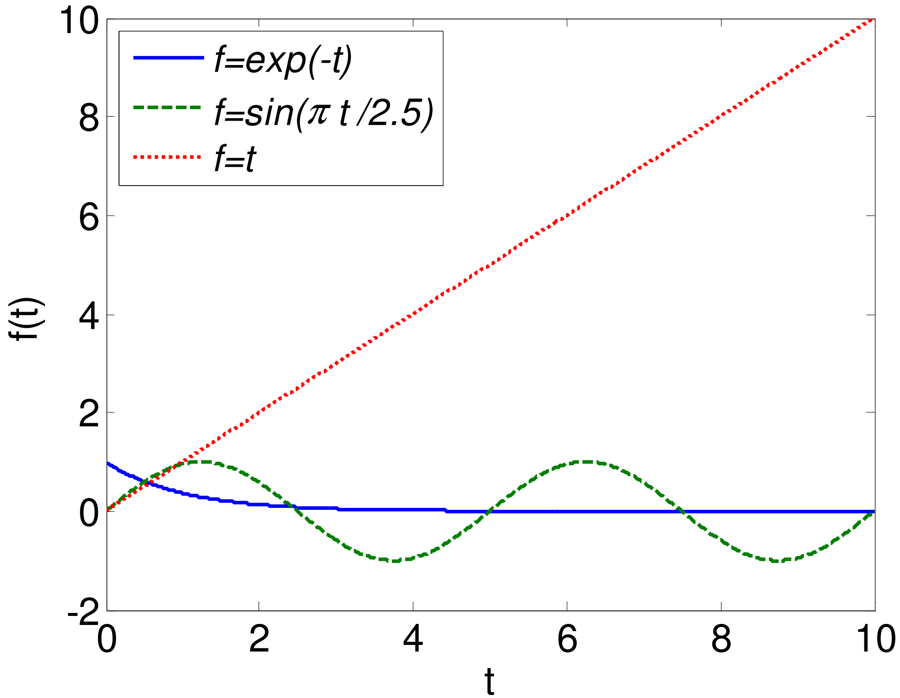
- standard criteria favor maximal time step
- predictions are even qualitatively in error

Discussion

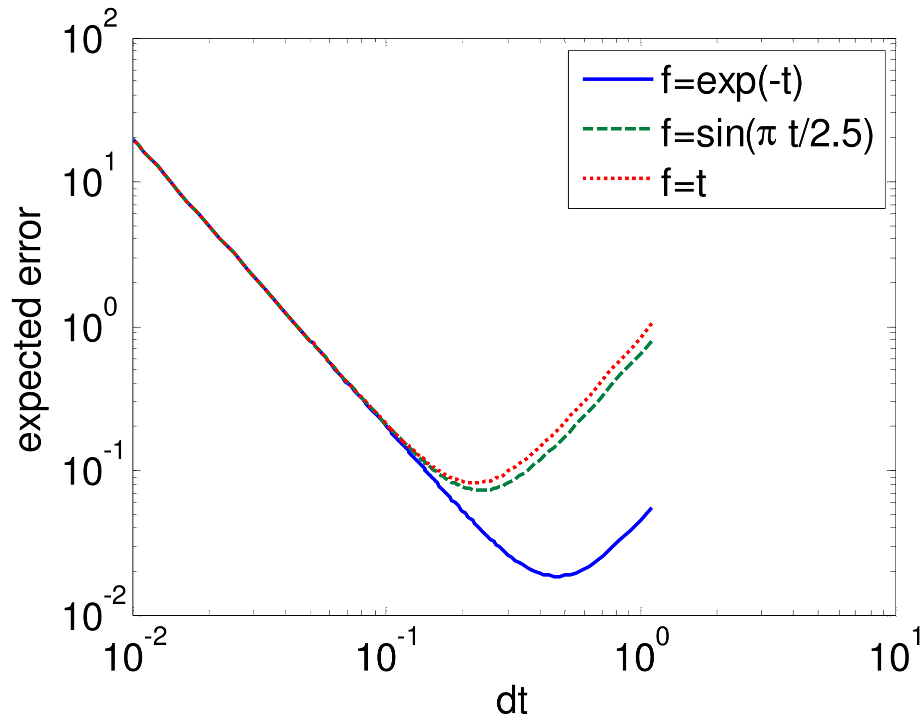
- *METER* criterion captures bias-variance trade-off
  crucial in the solution of ill-posed problems
- classical methods are even qualitatively wrong
- extension to nonlinear problems is straightforward
  by linearization and local analysis as in standard theory
- regularization parameter λ incorporates naturally
- main drawback: initial guess of unknown function f required

Dependence on Initial Guess

Test functions for initial guess



Resulting optimal time steps



- general trend independent of initial guess but property of kernel
- optimal time step halved ⇒ average deviation larger by factor 4
- dependency exists but is modest in the present example

Robust METER Designs

Assume: knowledge about function class
e.g., *f=exp(-kt)*

→ standard robust design strategies

Average design:

$$\min E_f \left[\text{Meter\_objective(f)}\right]$$

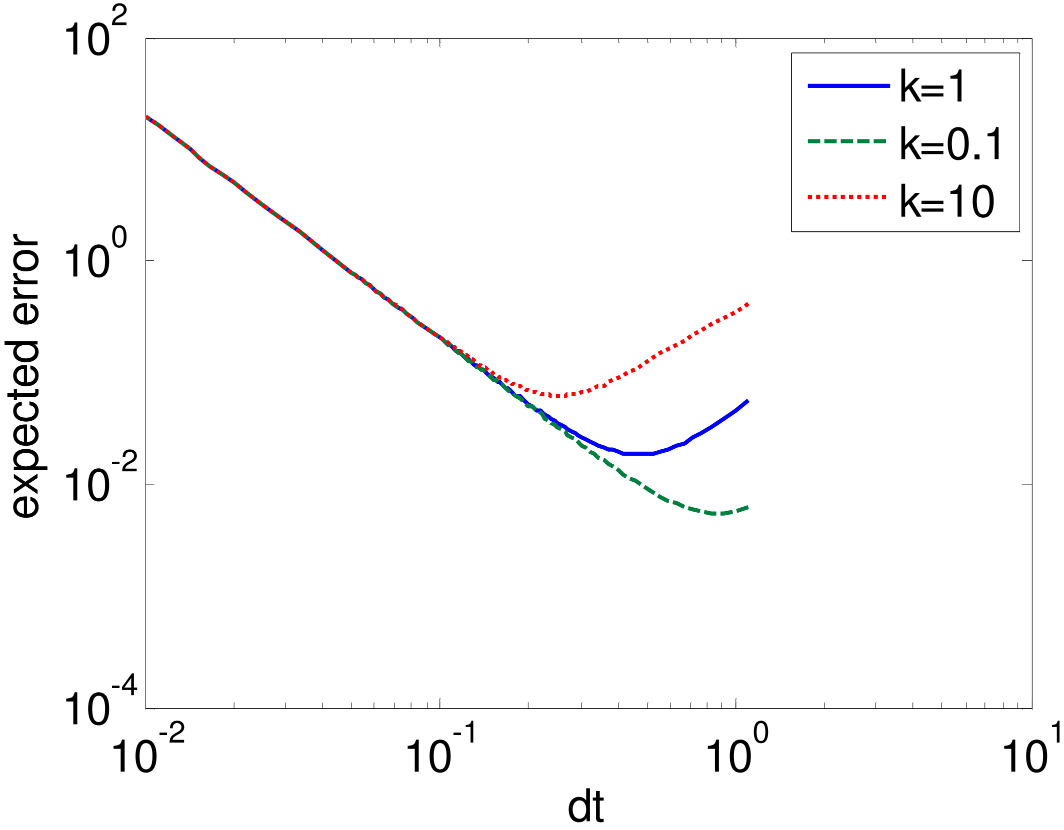
Worst-case design:

$$\min_d \max_f \left[\text{Meter\_objective(f)}\right]$$

TU Delft

Robust METER: Min-Max Design

Assume:
k=[0.1,10]

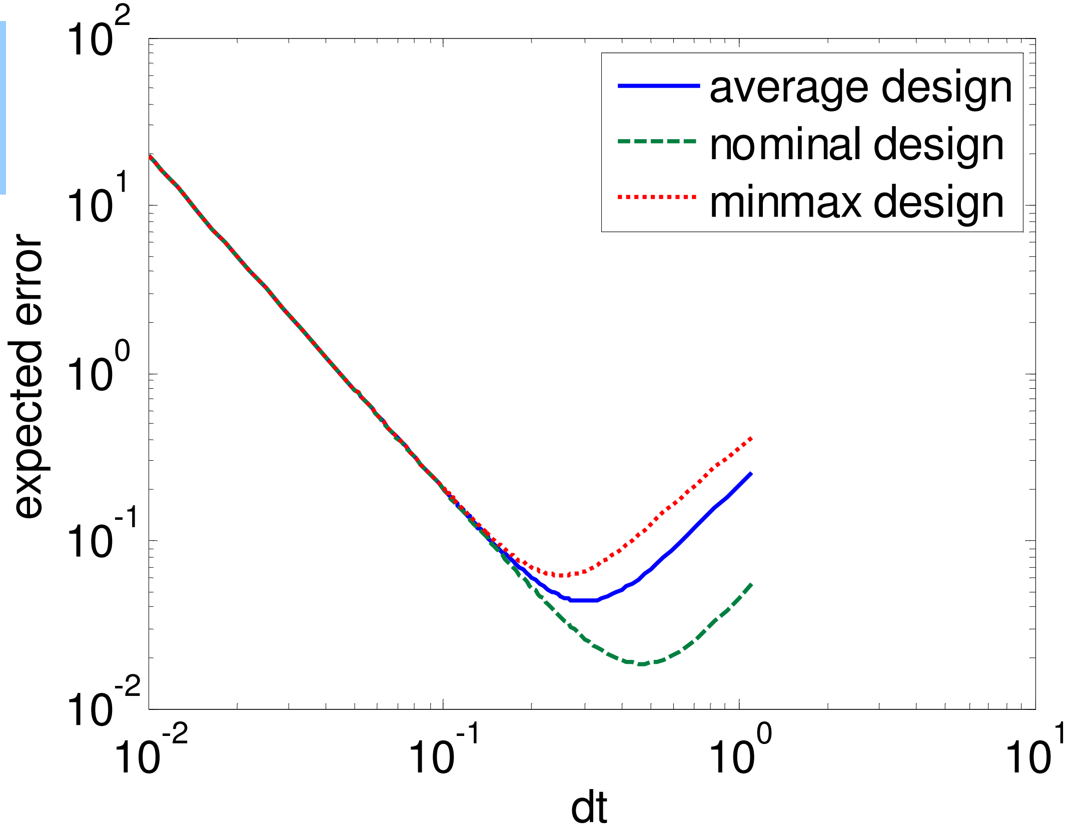


- worst case = largest decay rate
- simple solution of complex optimization problem
- enhanced problem understanding

TU Delft

Robust METER: Average Design

Assume:
k=[0.1,10]
equidistributed



- average design lies in-between
- strategies generalize to more complex cases

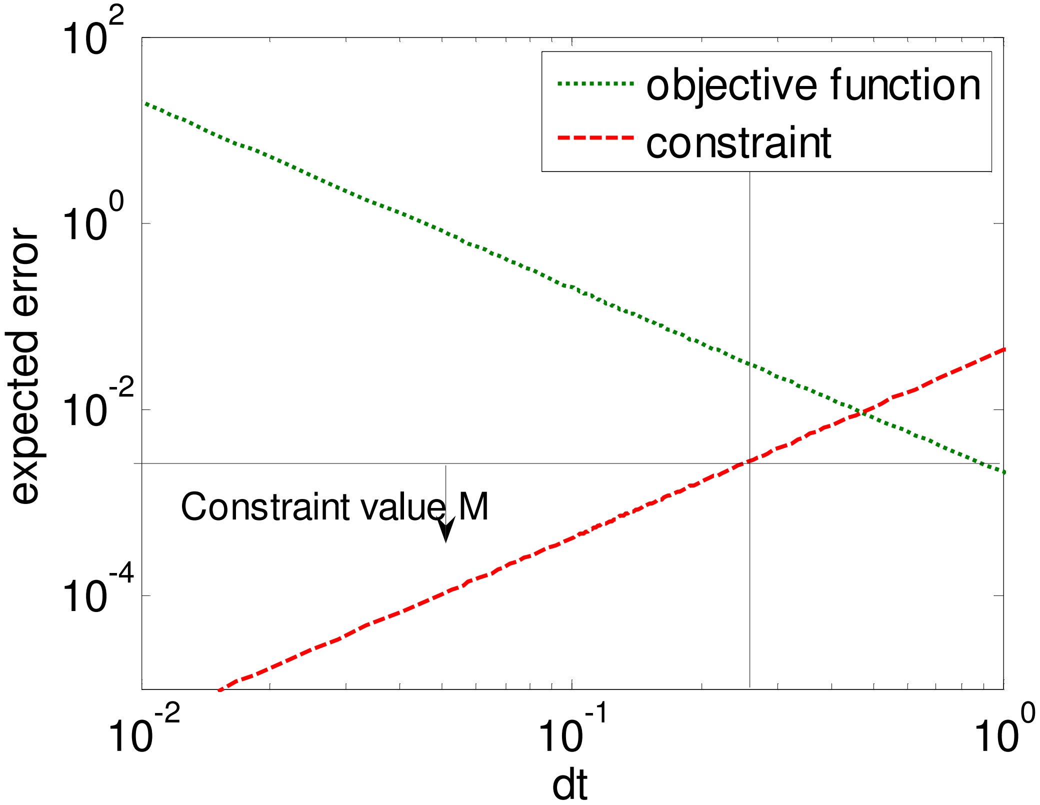
Constrained METER Design

Use bound for bias term as constraint

$$\min \operatorname{trace}\left[K^{+}(\lambda)\left(K^{+}(\lambda)\right)^{T}\right]$$

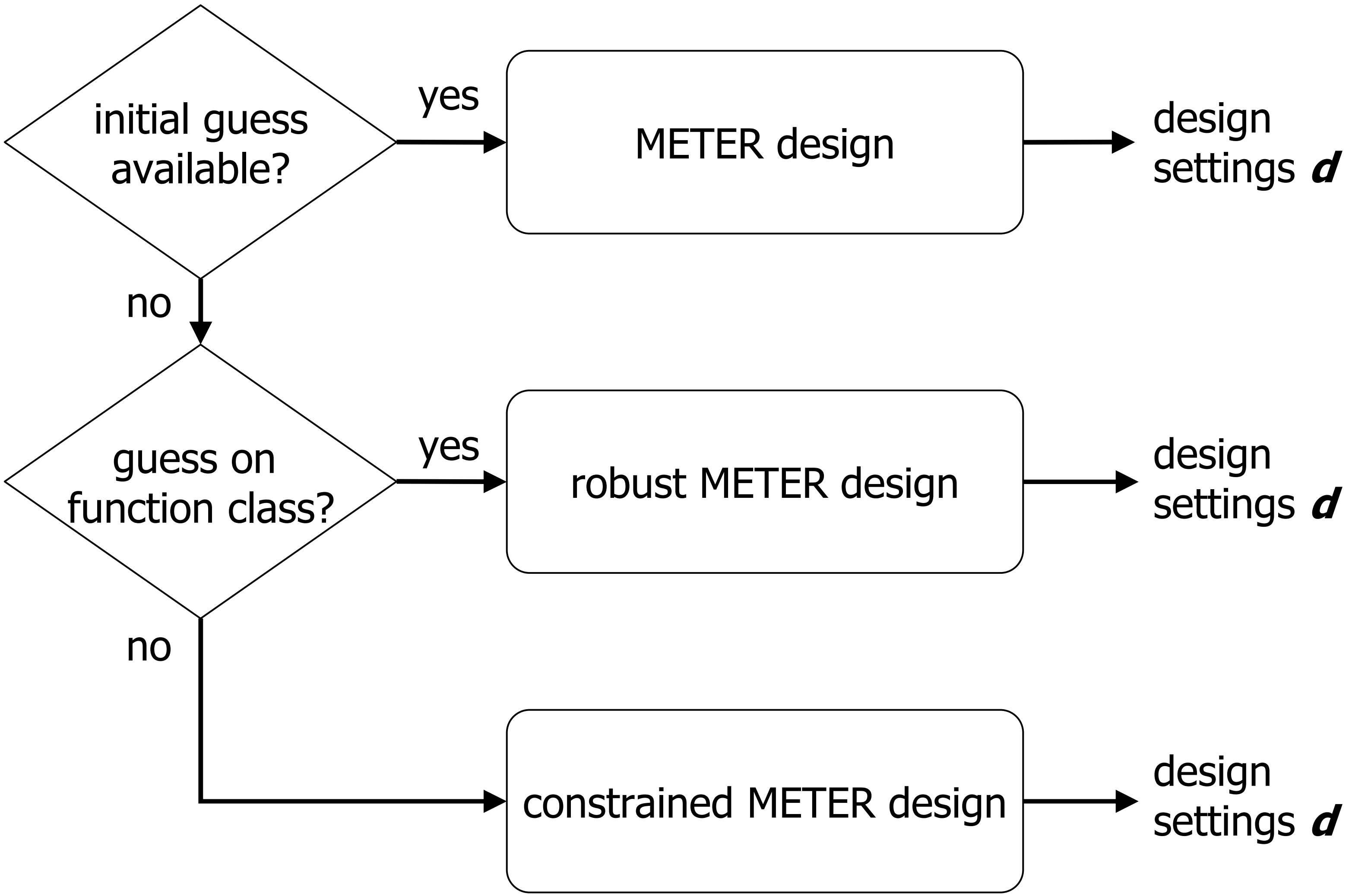
$$s.t. \quad \left\|\left(K^{T}K+\lambda L^{T}L\right)^{-1}\lambda L^{T}L\right\| \leq M$$

no knowledge about true function required

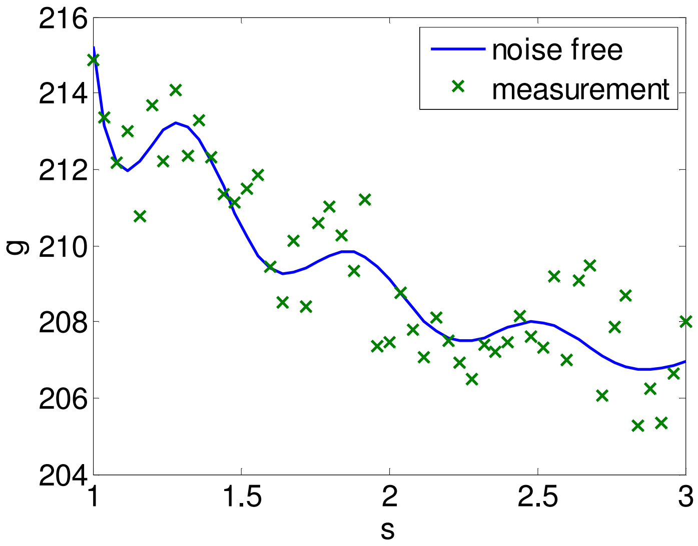


- directly applicable only to linear case
- requires guess on regularization parameter

Handling of prior knowledge



Particle sizing by light scattering



How to select measurement range *s* ?

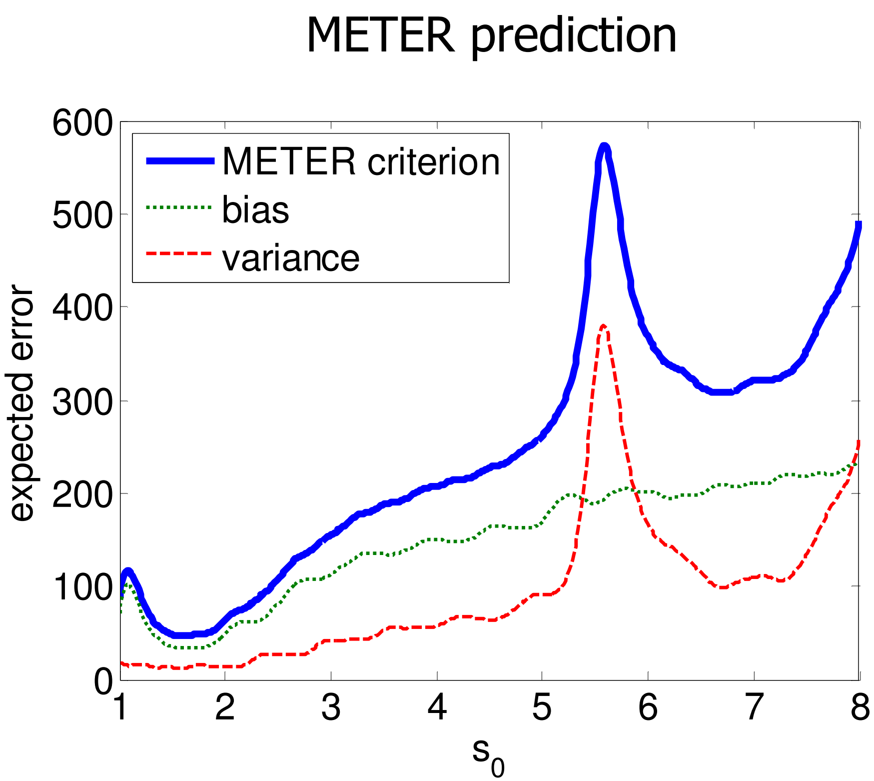
collected spectra

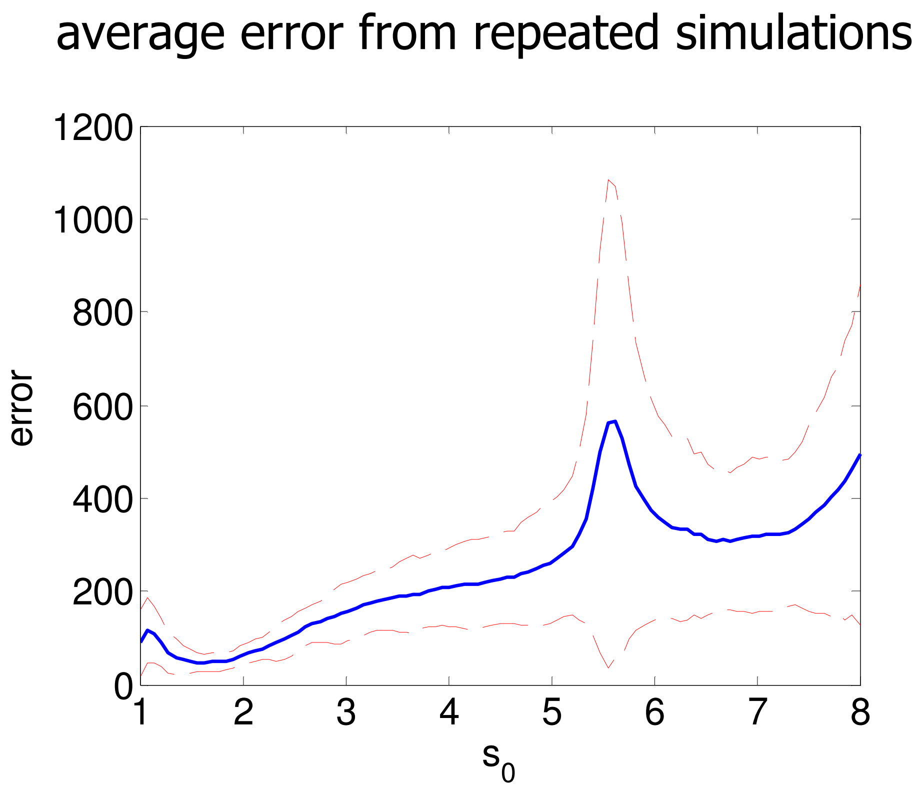
particle size number

$$g(s) = Kf(s) = \int_0^\infty \pi r^2 k(sr) N(r) dr$$

$$\text{with } k(x) = 2 - 4\frac{\sin(x)}{x} + 4\frac{(1-\cos(x))}{x^2}$$

METER design for light scattering

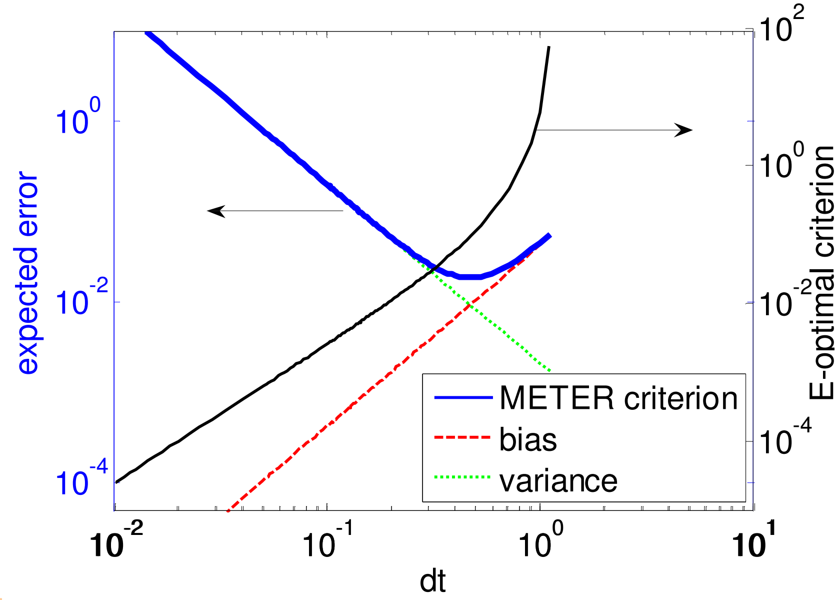




- highly nontrivial shape of bias and variance
- global trends similar in bias and variance but large differences in local details
- METER design captures true behavior

Summary: Inverse Problems

- *METER* captures bias-variance trade-off in ill-posed problems
- *METER* allows separate study of bias and variance contributions
- *METER* approach adaptable to level of *a priori* knowledge



Sound framework
for optimal experimental design
for ill-posed problems

Optimal experimental design

3. Design of Experiments

Of the two, design and analysis, the former is undoubtedly of greater importance. The damage of poor design is irreparable; no matter how ingenious the analysis, little information can be salvaged from poorly planned data. On the other hand, if the design is sound, then even quick and dirty methods of analysis can yield a great deal of information.

Box & Hunter (1965)

Watch out:
We are about to pimp your experiment!

Acknowledgements

- RWTH Aachen

  E. Kriesten, W. Marquardt (Lehrstuhl für Prozesstechnik)

  V. Göke, H.-J. Koß (Lehrstuhl für Technische Thermodynamik)

- ETH Zurich

  H.C. Öttinger (Polymer Physics)

- Deutsche Forschungsgemeinschaft

  for funding

References

- A. Bardow. Optimal experimental design of ill-posed problems: The METER approach. Accepted for: Computers & Chemical Engineering, 2007, doi:10.1016/j.compchemeng.2007.05.004
- A. Bardow, V. Göke, H.-J. Koß, W. Marquardt. Ternary diffusivities by model-based analysis of Raman spectroscopy measurements. AIChE Journal, 52(12): 4004-4015, 2006.
- A. Bardow, V. Göke, H.-J. Koß, K. Lucas, W. Marquardt. Concentration-dependent diffusion coefficients from a single experiment using model-based Raman spectroscopy. Fluid Phase Equilibria, 228-229: 357-366, 2005.
- A. Bardow, W. Marquardt. Incremental and simultaneous identification of reaction kinetics: methods and comparison. Chemical Engineering Science, 59(13): 2673-2684, 2004.
- A. Bardow, W. Marquardt. Identification of diffusive transport by means of an incremental approach. Computers & Chemical Engineering, 28(5): 585-595, 2004.
- A. Bardow, W. Marquardt, V. Göke, H.-J. Koß, K. Lucas. Model-based measurement of diffusion using Raman spectroscopy. AIChE Journal, 49(2): 323-334, 2003.