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 diffusion experiment

take an old car...
make it faster, nicer,... – better!

take an old experiment...
make it faster, nicer,... – better!
Model-based experimentation

Stage I: Model discrimination
- Propose model candidates
- Experiment design
- Experiment
- Data analysis
- Suitable model structure

Stage II: Parameter estimation
- Suitable model structure
- Experiment design
- Experiment
- Data analysis
- Optimal parameter values
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

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

Least-squares-type objective

Stage II: Parameter estimation

suitable model structure

experiment design

data analysis

experiment

optimal parameter values
Optimal Experimental Design: Idea

- set free variables $x$ such that information on parameter $\theta$ is maximized
- maximize curvature of parameter estimation objective

OED for parameter estimation:
\[
\max_x \frac{\partial^2 \phi}{\partial \theta^2}
\]
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 ($\Delta t = 10\text{ s}$, $\Delta z = 20\text{ µm}$)
- measurement error: statistical $\leq 0.2\text{ mol-\%}$, systematic $\leq 0.5\text{ mol-\%}$,

(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

[Diagram showing measurement zones and variables]
Optimal Experimental Conditions

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

- scaled objective $\zeta$-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 $\rightarrow$ optimal Fourier number
Wall effect in ternary system

Toluene-Dioxane-Cyclohexane; $x_{\text{Dioxane}} \approx \text{const}$

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 ($\phi^{(2)} = \phi^{(1)} + 90^\circ$)
- include constraints to ensure hydrodynamic stability
Ternary Raman experiment
Ternary Results

Diffusivities from a single Raman experiment

\[
\begin{align*}
D_{ij}^V [10^{-9} \text{ m}^2/\text{s}] \\
D_{11}^V &\quad D_{22}^V \\
D_{12}^V &\quad D_{21}^V
\end{align*}
\]

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

Diffusivities from two optimal Raman experiments

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

with reference data from Weingärtner et al. (1994)
Summary: Diffusion experiments

- 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

One stage: Inverse problem

- experimental data
- experiment design
- data analysis
- experiment
- model structure and parameters

Two stage approach:
- "Divide & conquer"
  - 😊 for \textit{a priori} knowledge
  - 😞 restricted search space

One stage approach:
- 😊 explore full search space
- 😞 no \textit{a priori} knowledge required
  → higher requirements on data
  → tougher estimation problem
Illustrating example: Batch reactor

\[ n(t) = \delta \sin(\omega t) \]

Estimation problem

\[
\begin{align*}
\min_r & \quad \|c - \tilde{c}\| \\
\text{s.t.} & \quad c(t) - c_0 = \int_0^t r(t) dt
\end{align*}
\]

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)

But...
- arbitrary large errors for bounded \( \delta \)
- errors governed by error level and frequency
- estimation errors do not depend continuously on data

Problem is ill-posed
Tikhonov regularization

Estimation problem

$$\min_{r} \|c - \tilde{c}\|$$

s.t. $c(t) - c_0 = \int_{0}^{t} r(t) dt$

$$g(t) = \int_{T} K(t, s; d) f(s) ds$$

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")

smoothness penalty

bias-variance trade-off crucial for success

error
data error
regularization error
regularization $\lambda$
Optimal Experimental Design: Idea

- set free variables \( \mathbf{x} \) 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

"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

\[
E\left(\left\| f - \hat{f} \right\|_{L^2}^2 \right) = \left\| f - K^+ (\lambda) Kf \right\|_{L^2}^2 + \sigma^2 \text{trace} \left( K^+ (\lambda) (K^+ (\lambda))^T \right) + \text{bias due to regularization} + \text{variance in estimate}
\]

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 $c(t)$
  \[
  c(t) = \int_{0}^{t} r(t') dt' \quad \Rightarrow \quad r(t) = \frac{dc(t)}{dt}
  \]

- 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 $\Delta t$ too large \(\rightarrow\) approximation error
- Time step $\Delta t$ 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
Comparison with Numerical Computations

- predicted behavior is found in practice
- small difference from approximate bias term
\[ r(t) = \exp(-t) \]
\[ t = [0,10] \]

Comparison with Classic Criteria

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

\( E \)-optimal design: max min eigenvalue of Fisher matrix
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 $\lambda$ incorporates naturally

• main drawback: initial guess of unknown function $f$ required
Dependence on Initial Guess

Test functions for initial guess

- $f(t) = e^{-t}$
- $f(t) = \sin\left(\frac{\pi t}{2.5}\right)$
- $f(t) = t$

Resulting optimal time steps

- general trend independent of initial guess but property of kernel
- optimal time step halved $\Rightarrow$ 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:

$$\min E_f [\text{Meter_objective}(f)]$$

Worst-case design:

$$\min \max_d \max_f [\text{Meter_objective}(f)]$$
Robust METER: Min-Max Design

Assume: $k=[0.1,10]$

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

\[
\min \text{trace} \left[ K^+ (\lambda)(K^+ (\lambda))^T \right] \\
\text{s.t. } \left\| (K^T K + \lambda L^T L)^{-1} \lambda L^T L \right\| \leq M
\]

- no knowledge about true function required

- directly applicable only to linear case
- requires guess on regularization parameter
Handling of prior knowledge

- Initial guess available?
  - Yes: METER design
    - Design settings $d$
  - No: no

- Guess on function class?
  - Yes: robust METER design
    - Design settings $d$
  - No: constrained METER design
    - Design settings $d$
Particle sizing by light scattering

How to select measurement range \( s \) ?

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

with \( k(x) = 2 - 4 \frac{\sin(x)}{x} + 4 \frac{(1 - \cos(x))}{x^2} \)
METER design for light scattering

METER prediction

average error from repeated simulations

- 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 \textit{a priori} knowledge

Sound framework for optimal experimental design for ill-posed problems
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)
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


