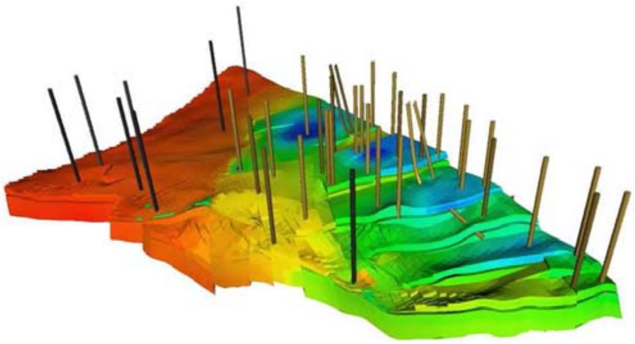


Parameter Estimation and Identifiability in Large-Scale Reservoir Models

Paul M.J. Van den Hof

co-authors: Jan Dirk Jansen, Gijs van Essen, Jorn Van Doren, Okko Bosgra, Sippe Douma, Amin Rezapour

THE VALUE OF SMARTNESS

ISAPP

Seminar Stanford University, Palo Alto, CA, 19 January 2011

Delft University of Technology

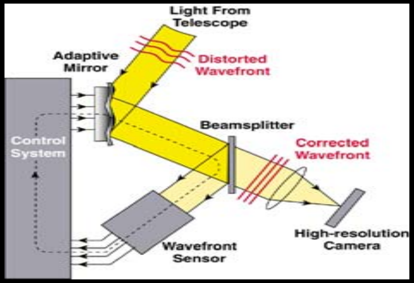
Oldest and largest of 3 Technical Universities in the Netherlands:
Delft – Eindhoven - Twente

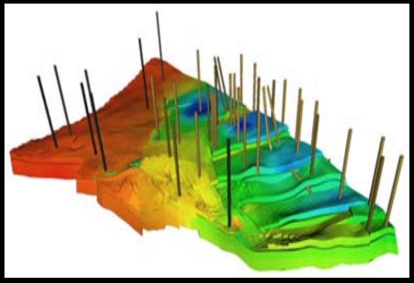
Founded in 1842 as an engineering school

Now: 5000 employees, of which 2700 scientists, 15,000 students, distributed over 8 engineering faculties:

- Electrical, Math, Comp. Science
- Aerospace Engineering
- Applied Sciences
- Mechanical, Maritime, Mat. Eng.
- Civil Engin., Earth Sciences
- Industrial Design
- Techn. Policy Making and Manag
- Architecture

Delft Center for Systems and Control





University wide center and department within Mechanical Engineering

- 16 scientific (tenured) staff
- 40 PhD students
- 15 Postdocs
- 40 MSc students / year

- BSc programs ME, EE, AP,
- MSC programs ME, EE, AP
- 3TU MSc Systems and Control
- PhD programme DISC

Fundamentals:

•Modelling, control and optimization of complex, non-linear and hybrid systems
•Signal analysis, signal processing and data-based modelling (identification for control)

Mechatronics and Microsystems:

- Automotive systems
- Microfactory
- AFM nano-positioning
- Smart optics systems
- Electron microscopes
- Robotics

Traffic and Transportation:

•Discrete event and hybrid systems
•Distributed multi-agent systems
•Optimal adaptive traffic control
•Advanced driver assistance systems

Sustainable Industrial Processes:

- Increase of scale in process operation
- More flexibility in operation
- Economic optimization under operating constraints
- Process intensification
- Towards model-based process management
- hydrocarbon reservoir optimization; crystallization

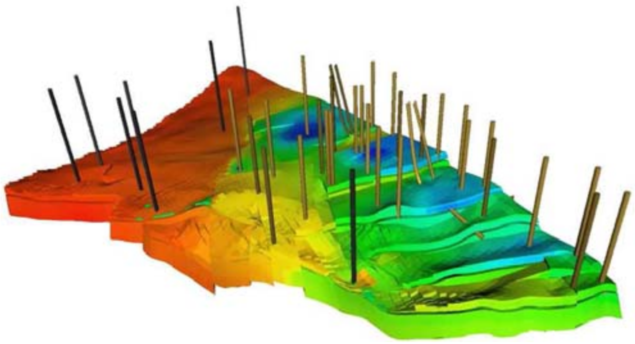
Model-Based Process Control and Optimization

Realize economic efficient operation of dynamic processes through smart (knowledge intensive) operation strategies

Smart = Model-Based, collecting all relevant information of the process, i.e. physical knowledge as well as experimental data

Applications

- Water flooding in reservoir engineering
- Crystallization processes
- Industrial hydrocarbon/gas-to-liquid production processes
- Intensified reaction systems (process intensification)
- Multiphase-flow systems (ILS)



Parameter Estimation and Identifiability in Large-Scale Reservoir Models

Paul M.J. Van den Hof

co-authors: Jan Dirk Jansen, Gijs van Essen, Jorn Van Doren, Okko Bosgra, Sippe Douma, Amin Rezapour

THE VALUE OF SMARTNESS

ISAPP

Seminar Stanford University, Palo Alto, CA, 19 January 2011

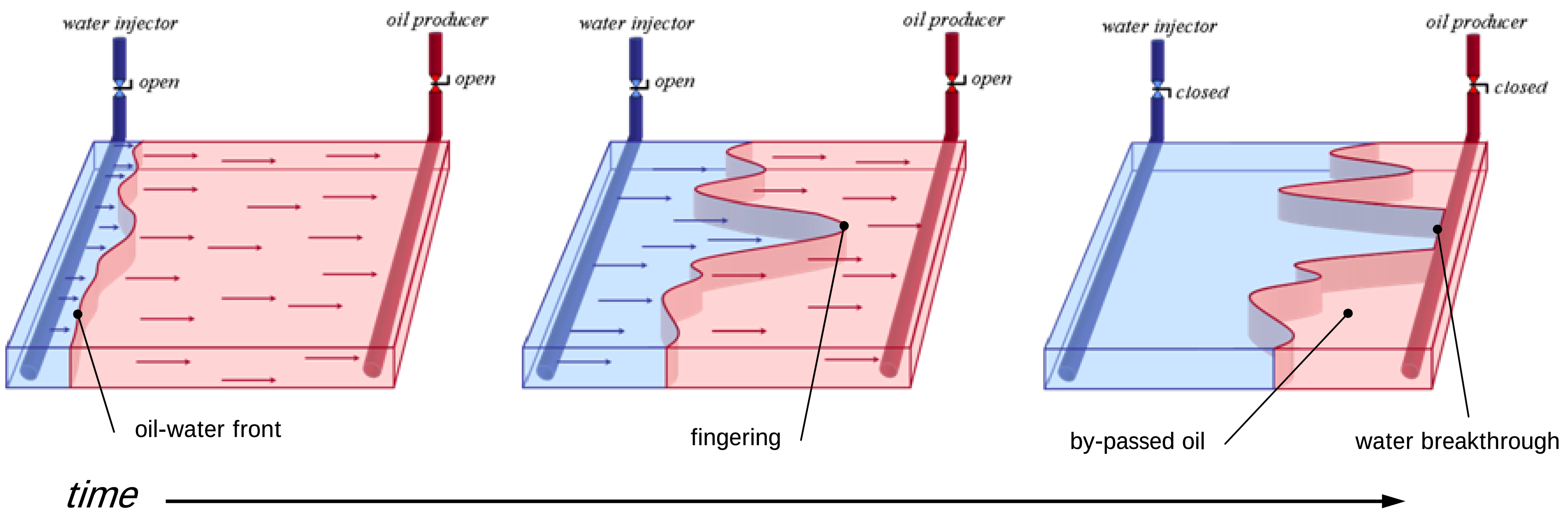
Contents

- Introduction – closed-loop reservoir management
- Model estimation
- Identifiability and model structure approximation
- Simple example
- Extension to data driven approaches: two-stage procedure
- Summary

Introduction

Water flooding

- Involves the injection of water through the use of injection wells
- Goal is to increase reservoir pressure and displace oil by water
- Production to be optimized by manipulating injector and producer valves over life-cycle

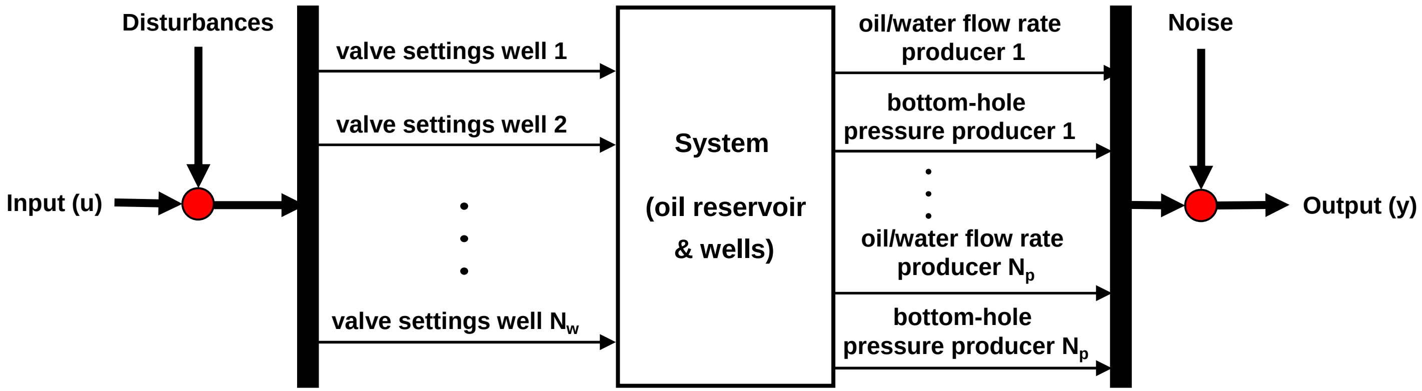


[D.R. Brouwer, 2004]

The Models

System involves the reservoir, wells and sometimes surface facilities

- Inputs: control valve settings of the wells (injectors and producers)
  - Smart wells: multiple (subsurface) valves
- Outputs: (fractional) flow rates and/or bottomhole pressures
  - Smart wells: multiple (subsurface) measurement devices



Governing differential equations
isothermal two-phase (oil-water) flow

Mass balance:

$$\nabla(\rho_i u_i) + \frac{\partial}{\partial t}(\phi \rho_i S_i) = 0 \qquad i = \{o, w\}$$

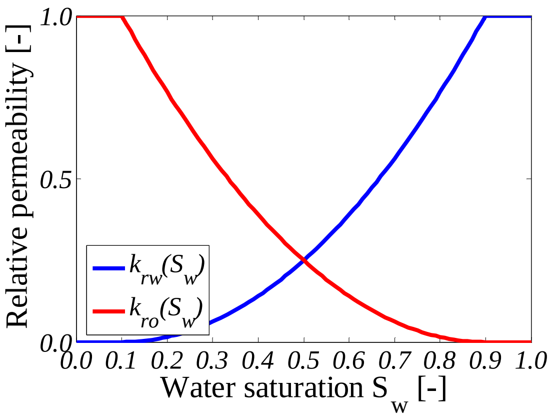
Momentum (Darcy's law):

$$u_i = -k \frac{k_{ri}}{\mu_i} \nabla p_i \qquad i = \{o, w\}$$

Variables: p_o, p_w, S_o, S_w

Saturations satisfy: $S_o + S_w = 1$

Simplifying assumptions, a.o.: $p_o = p_w$



Discretization in space and time

State space model:

$$
\begin{aligned}
V(x_t)\dot{x}_t &= T(x_t)x_t + q_t; \qquad x_0 \\
y_t &= h(x_t)
\end{aligned}
$$

$$
\begin{aligned}
y^T &= [p_{well}^T \; q_{well,o}^T \; q_{well,w}^T] \\
x^T &= [p_o^T \; S_w^T]
\end{aligned}
$$

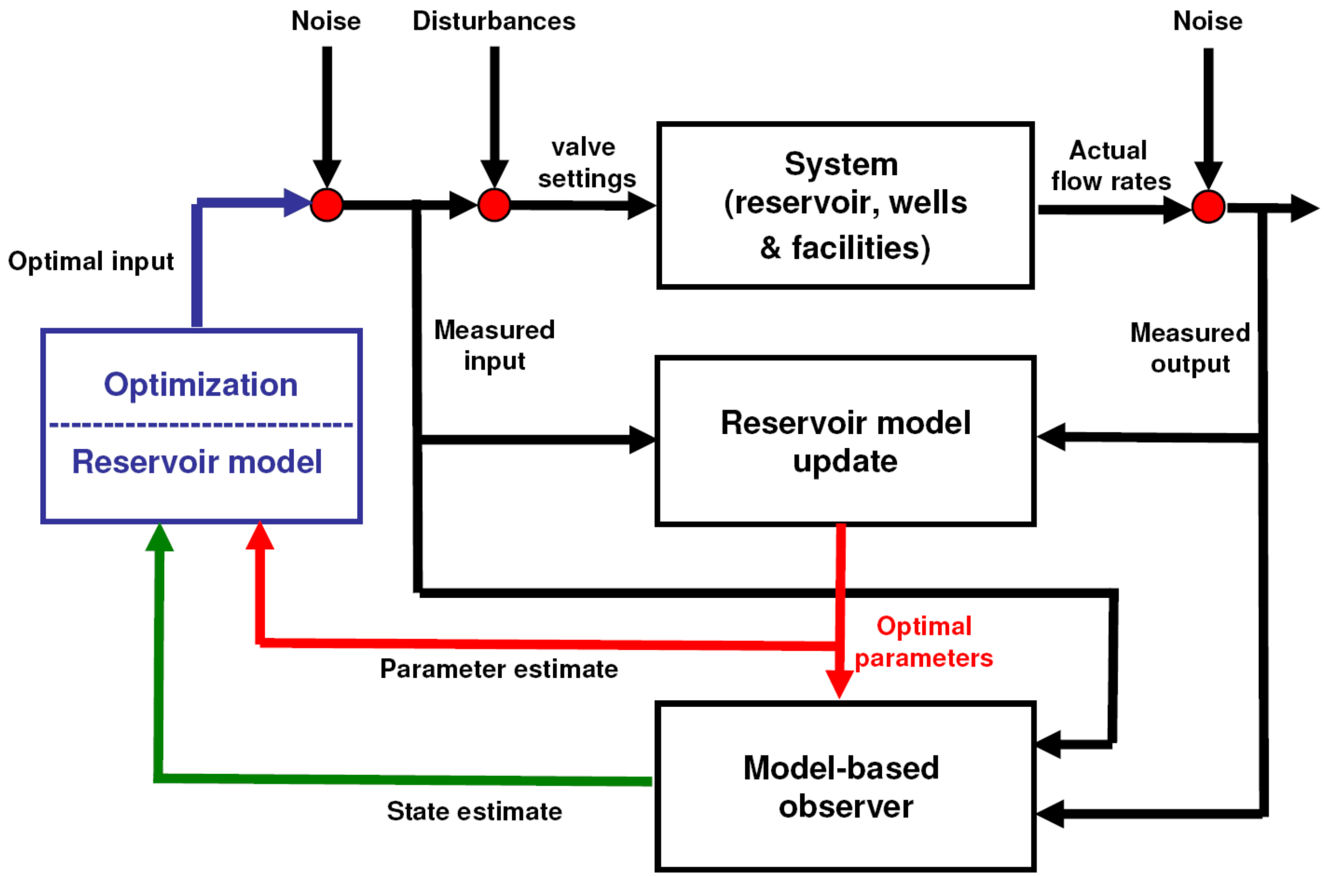
After discretization in space (and time):

$$
g(x_{k+1}, x_k, u_k, \theta) = 0 \qquad dim(x) \approx 10^4 - 10^6
$$

$$
y_k = h(x_k)
$$

and θ typically the permeabilities in each grid block

Closed-loop Reservoir Management



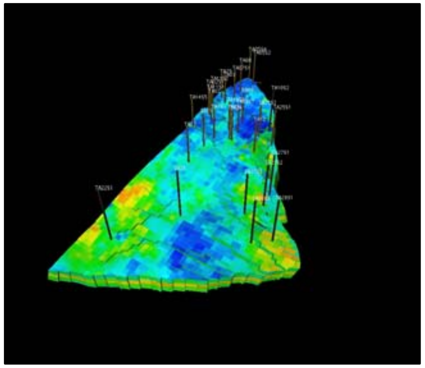
Closed-loop strategy:

- Natural and standard strategy in process technology (feedback strategy) and elsewhere
- Allows to reduce sensitivity with respect to uncertainties in the data (disturbances), and uncertainties in the model (robustness)
- Through the on-line use of the model, generally favouring models of limited complexity
- For models with linear dynamics, the understanding of the control-relevant parts of the models is well developed

Contents

- Introduction – closed-loop reservoir management
- Model estimation
- Identifiability and model structure approximation
- Simple example
- Extension to data driven approaches: two-stage procedure
- Summary

Model estimation



Is there a problem with the online use of highly complex first-principles models?

Not if:

- Computation time is not a problem, and
- They exactly describe reality

In case of model uncertainty, one would like to use measurement data to retrieve information.

Model estimation

Estimating models from experimental data can be done in different ways:

1. Estimating the dynamics between input and output in a generic (black box) model structure, using as few parameters as possible (to be determined by the data)

   Relatively “easy” for linear dynamics;

   Harder for nonlinear reservoirs
   (nonlinear behaviour dependent on front-location, single batch process, experimental limitations)

Model estimation

Estimating models from experimental data can be done in different ways:

2. Estimating the parameters (permeabilities) in a physics-based model structure possible using prior information on their numerical values

number of parameters has to be observed → identifiability

Contents

Identifiability

- Consider nonlinear model structure $\hat{\mathbf{y}} = \mathbf{h}(\boldsymbol{\theta}, \mathbf{u}; \mathbf{x}_0)$
  with $\hat{\mathbf{y}}$ being a prediction of the measured outputs $\mathbf{y} := [y_1^T \cdots y_N^T]^T$

The model strucure is

- Locally identifiable in $\boldsymbol{\theta}_m$ for given $\mathbf{u}$ and $\mathbf{x}_0$ if in neighbourhood of $\boldsymbol{\theta}_m$:

$$\{\mathbf{h}(\mathbf{u}, \boldsymbol{\theta}_1; \mathbf{x}_0) = \mathbf{h}(\mathbf{u}, \boldsymbol{\theta}_2; \mathbf{x}_0)\} \Rightarrow \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$$

[Grewal and Glover 1976]

Global properties are generally very hard to analyze (nonlinear)

- Notion of *identifiability* is instrumental in analyzing model structure properties
- It determines whether it is feasible at all to relate unique values to the physical parameter variables, on the basis of measured data

Testing local identifiability in model estimation

- Consider quadratic identification criterion based on prediction errors

$$V(\boldsymbol{\theta}) := \frac{1}{2}\boldsymbol{\epsilon}(\boldsymbol{\theta})^T \mathbf{P}_v^{-1}\boldsymbol{\epsilon}(\boldsymbol{\theta}), \quad \boldsymbol{\epsilon}(\boldsymbol{\theta}) = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{h}(\boldsymbol{\theta}, \mathbf{u}; \mathbf{x}_0),$$

- Hessian given by

$$\frac{\partial^2 V(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} = \frac{\partial \hat{\mathbf{y}}^T}{\partial \boldsymbol{\theta}} \mathbf{P}_v^{-1} \left(\frac{\partial \hat{\mathbf{y}}^T}{\partial \boldsymbol{\theta}} \right)^T + \mathbf{S}$$

- Local identifiability test in $\hat{\boldsymbol{\theta}} = \arg\min V(\boldsymbol{\theta})$: Hessian > 0

- With quadratic approximation of cost function around $\hat{\boldsymbol{\theta}}$: Hessian given by $\frac{\partial \hat{\mathbf{y}}^T}{\partial \boldsymbol{\theta}} \mathbf{P}_v^{-1} \left(\frac{\partial \hat{\mathbf{y}}^T}{\partial \boldsymbol{\theta}} \right)^T$

Testing local identifiability in identification

- Rank test on Hessian through SVD

$$\frac{\partial \hat{\mathbf{y}}^T}{\partial \boldsymbol{\theta}} \mathbf{P}_v^{-\frac{1}{2}} \bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}$$

- If $\boldsymbol{\Sigma}_2 = 0$ then lack of local identifiability

- SVD can be used to reparameterize the model structure through

$$\boldsymbol{\theta} = \mathbf{U}_1 \boldsymbol{\rho}, \quad \dim(\boldsymbol{\rho}) << \dim(\underline{\theta})$$

  in order to achieve local identifiability in $\boldsymbol{\rho}$

- Columns of $\mathbf{U}_1$ are basis functions of the identifiable parameter space

Testing local identifiability in identification

$$\left.\frac{\partial \hat{\mathbf{y}}^T}{\partial \boldsymbol{\theta}} \mathbf{P}_v^{-\frac{1}{2}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}$$

- What if $\boldsymbol{\Sigma}_2 \neq 0$ but contains (many) small singular values ?

  No lack of identifiability, but possibly very poor variance properties

- Identifiability mostly considered in a yes/no setting: qualitative rather than quantitative [Bellman and Åström (1970), Grewal and Glover (1976)]

- Approach: *quantitative* analysis of appropriate parameter space, maintaining physical parameter interpretation

Model structure approximation

- How to reduce the model structure in terms of its ***parameter space?***
  (different from "classical" model reduction, in which the model dynamics of a single model is reduced)

- **Objective:** obtain a physical parametrization (model structure) in which the parameters can be reliably estimated / validated from data.

Approximating the identifiable parameter space

Asymptotic variance analysis: $\mathrm{cov}(\hat{\boldsymbol{\theta}}) = J^{-1} = \left(\mathbb{E}\left[\left. \frac{\partial^2 V(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right|_{\hat{\boldsymbol{\theta}}} \right] \right)^{-1}$

with J = Fisher Information Matrix.

- Sample estimate of parameter variance, on the basis of $V(\boldsymbol{\theta})$:

$$cov(\hat{\boldsymbol{\theta}}) = \begin{cases} \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_1^{-2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2^{-2} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} & \text{for } \boldsymbol{\Sigma}_2 > 0 \\ \infty & \text{for } \boldsymbol{\Sigma}_2 = 0 \end{cases}$$

$$cov(\boldsymbol{U_1}\hat{\boldsymbol{\rho}}) = \mathbf{U_1}\boldsymbol{\Sigma}_1^{-2}\mathbf{U_1}^T$$

$$cov(\hat{\boldsymbol{\theta}}) > cov(\boldsymbol{U_1}\hat{\boldsymbol{\rho}}) \qquad \text{if } \boldsymbol{\Sigma}_2 > 0$$

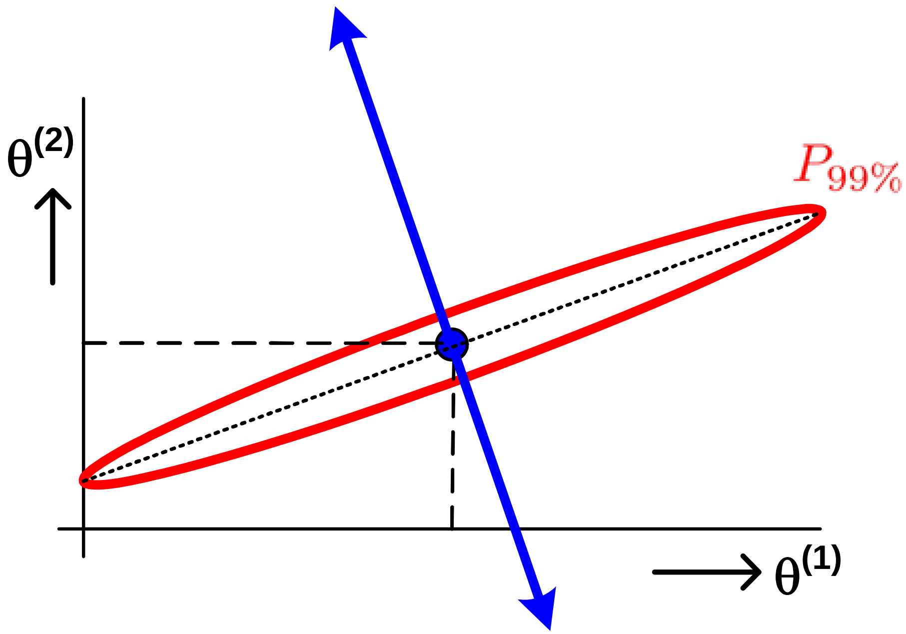
Approximating the identifiable parameter space

$$cov(\hat{\boldsymbol{\theta}}) > cov(\boldsymbol{U}_1\hat{\boldsymbol{\rho}}) \qquad \text{if } \boldsymbol{\Sigma}_2 > 0$$

- Discarding singular values that are small, reduces the variance of the resulting parameter estimate
- Particularly important in situations of (very) large numbers of small s.v.'s
- Model structure approximation (local)
- Quantified notion of identifiability – related to parameter variance

Approximating the identifiable parameter space

- Interpretation:
  Remove the parameter directions that are poorly identifiable (have large variance)



- This is different from removing the (separate) parameters for which the value 0 lies within the confidence bound [Hjalmarsson, 2005]

A Bayesian approach

How does this work out in estimation through Kalman-filter type algorithms (like EnKF)?

- State vector is augmented with unknown parameters and estimated simultaneously in a recursive algorithm

These algorithms can be interpreted in a Bayesian framework

A Bayesian approach

- Often applied method for dealing with overdetermination in parameter space:
- Incorporate prior knowledge term (regularization) in cost function

$$V_p(\boldsymbol{\theta}) := V(\boldsymbol{\theta}) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_p)\mathbf{P}_{\boldsymbol{\theta}_p}^{-1}(\boldsymbol{\theta} - \boldsymbol{\theta}_p)$$

where $\boldsymbol{\theta}_p$ is the prior parameter vector (with covariance $\mathbf{P}_{\boldsymbol{\theta}_p}$).

- When model output approximated with first-order Taylor expansion, then Hessian is

$$\frac{\partial^2 V_p(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} = \frac{\partial \mathbf{h}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} \mathbf{P}_v^{-1} \left(\frac{\partial \mathbf{h}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} \right)^T + \mathbf{P}_{\boldsymbol{\theta}_p}^{-1}$$

- “Always” identifiable, since $\mathbf{P}_{\boldsymbol{\theta}_p}$ full rank by construction!!

A Bayesian approach

Implications

$$V_p(\boldsymbol{\theta}) := \underbrace{V(\boldsymbol{\theta})}_{\text{data}} + \underbrace{\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}_p)\mathbf{P}_{\boldsymbol{\theta}_p}^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}_p)}_{\text{priors}}$$

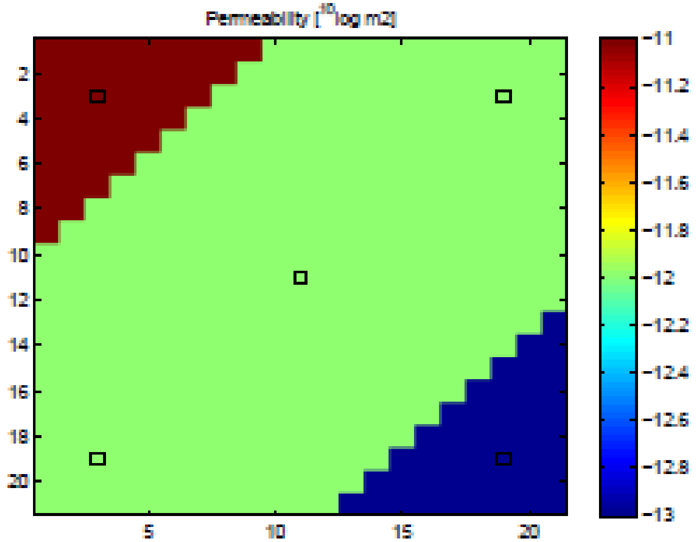
- Bayesian methods seem not to suffer from identifiability problems.......
- This includes all (extended) Kalman filter type algorithms. Where parameters are recursively estimated by augmenting the states
- Unique parameter estimates usually result, but
- In the parameter subspace that is poorly identifiable, estimated parameters will be heavily dominated by the prior information.
- Analysis of $V(\theta)$ can show identifiable directions (locally)

Simple reservoir example

[Van Doren, 2010]

2D two-phase example

(top view)

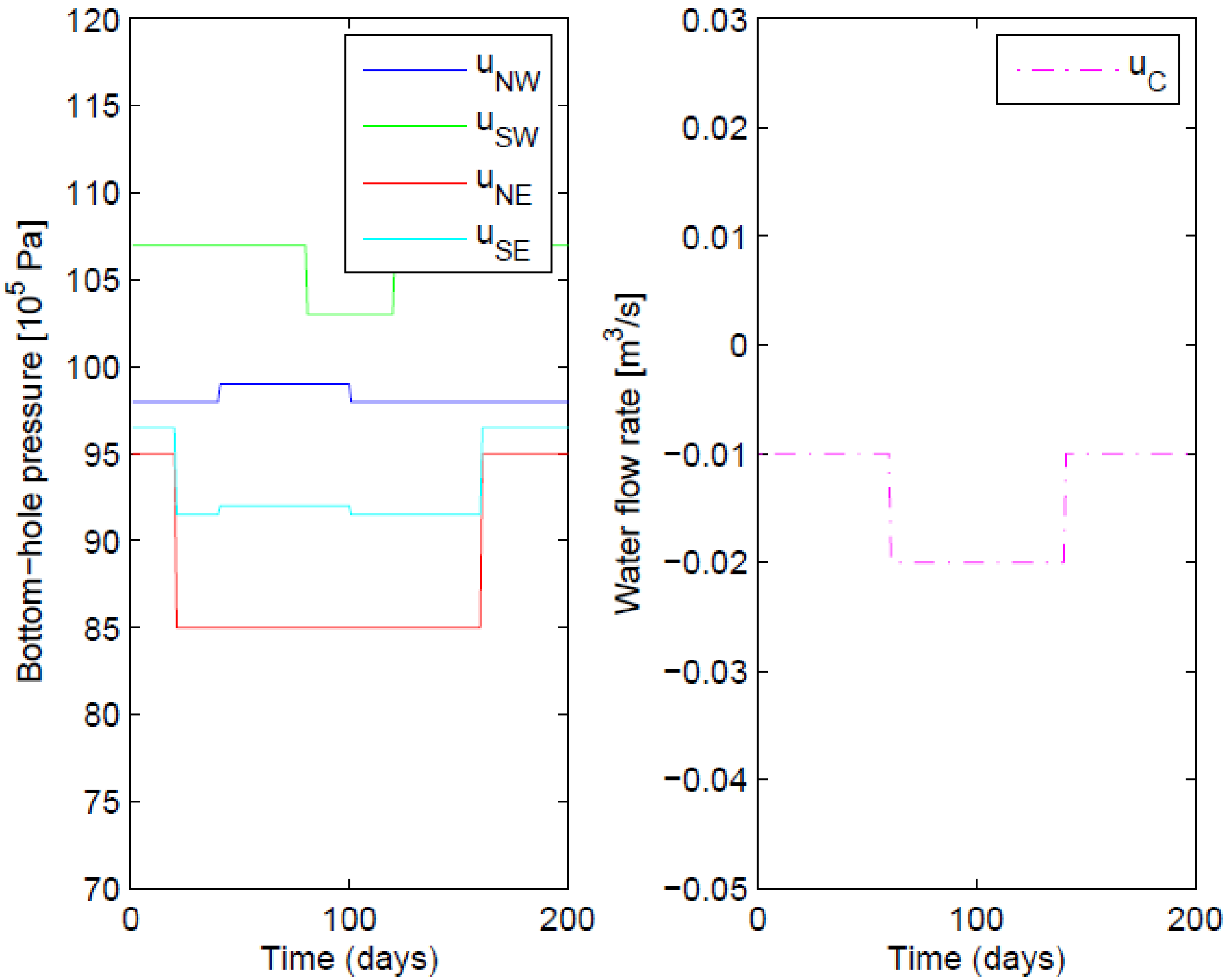


21 x 21 grid block permeabilities
5 wells; 3 permeability strokes

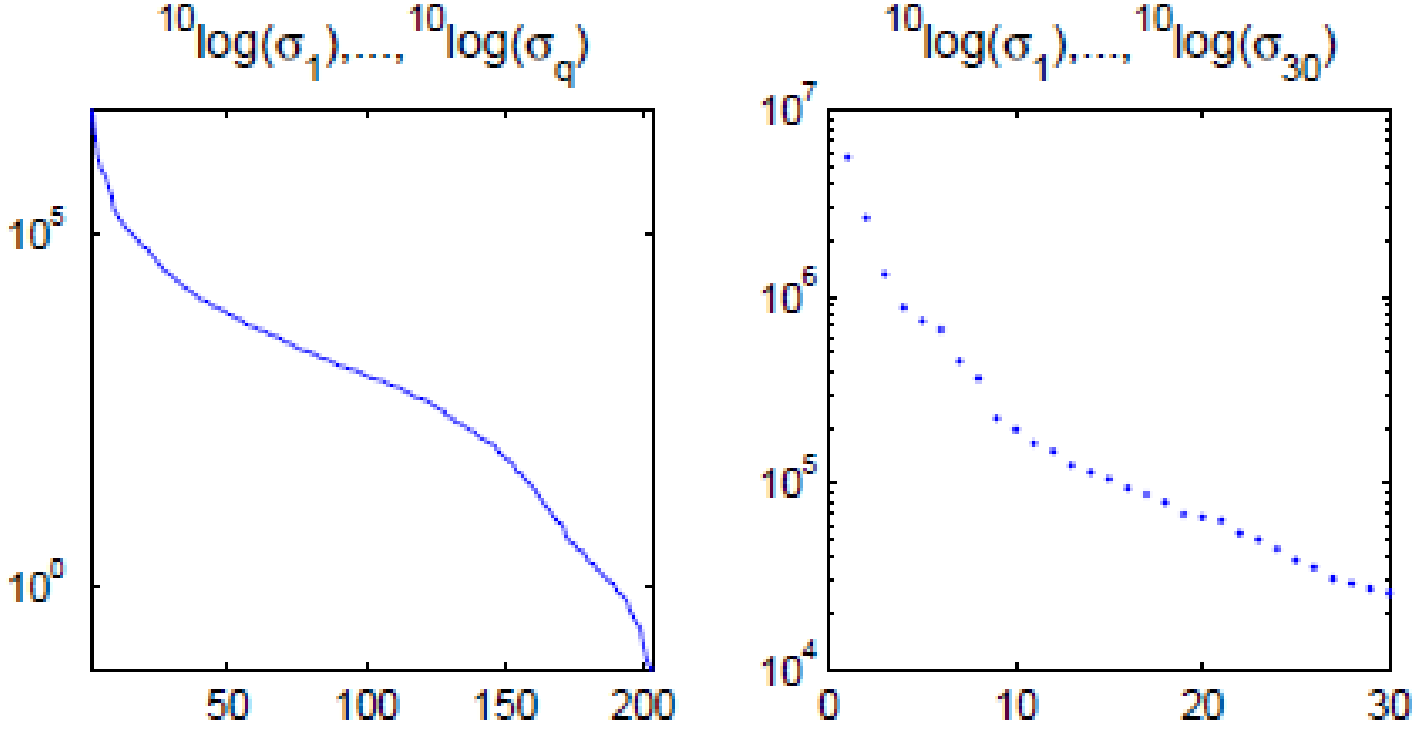
1 injector (centre)
4 producers (corners)

5 inputs: 1 injector flow-rate, and 4 bottom hole pressures
8 outputs: producer flow rates (water and oil)

Input excitation signals



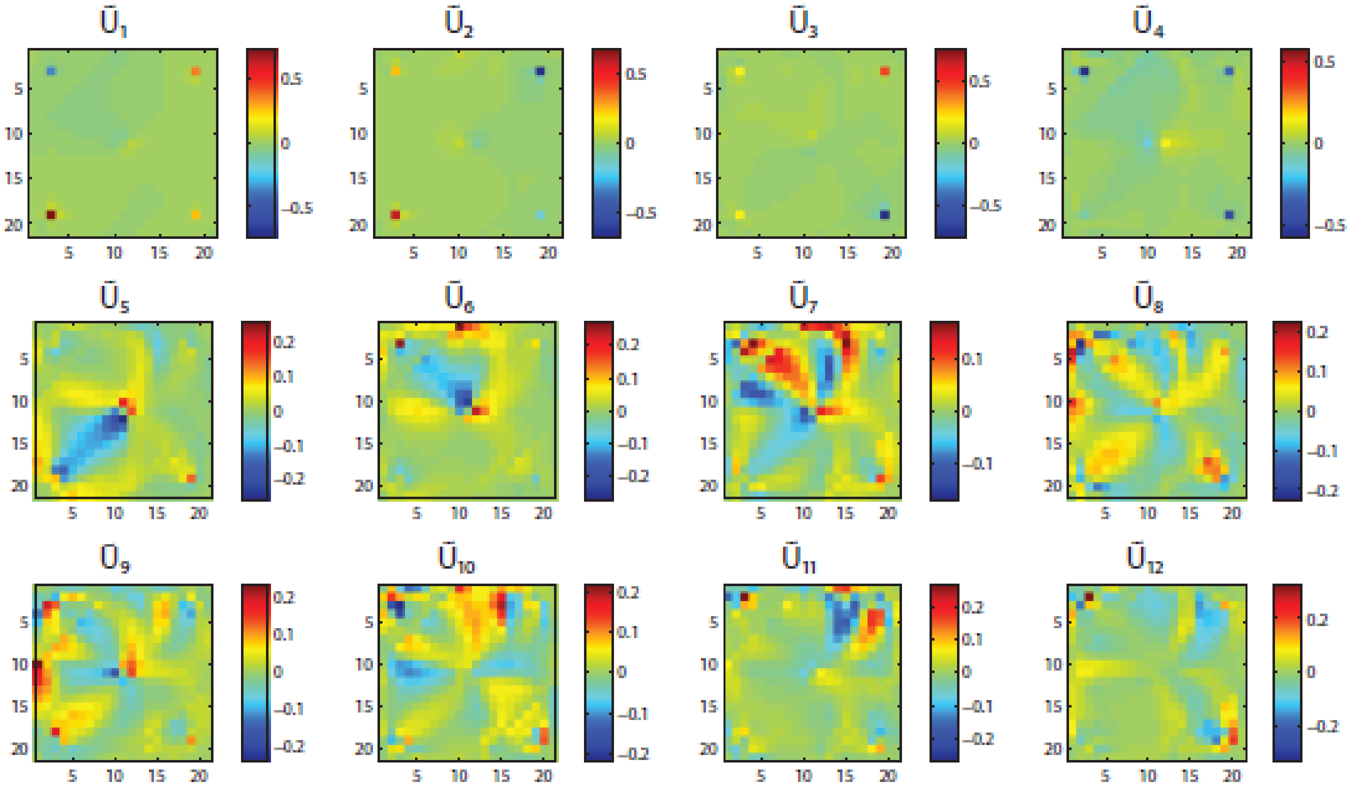
Simple reservoir example



All singular values (left) and first 30 (right)

Simple reservoir example

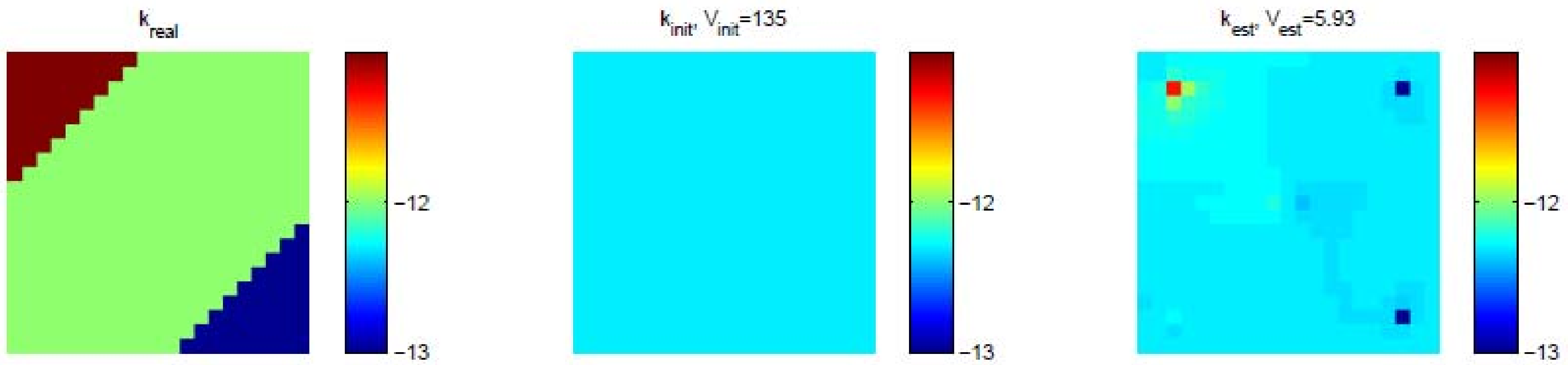
Singular vectors can be projected on the grid:



First 12 singular vectors

Simple reservoir example

Using the reduced parameter space –iteratively- in estimation:



exact field

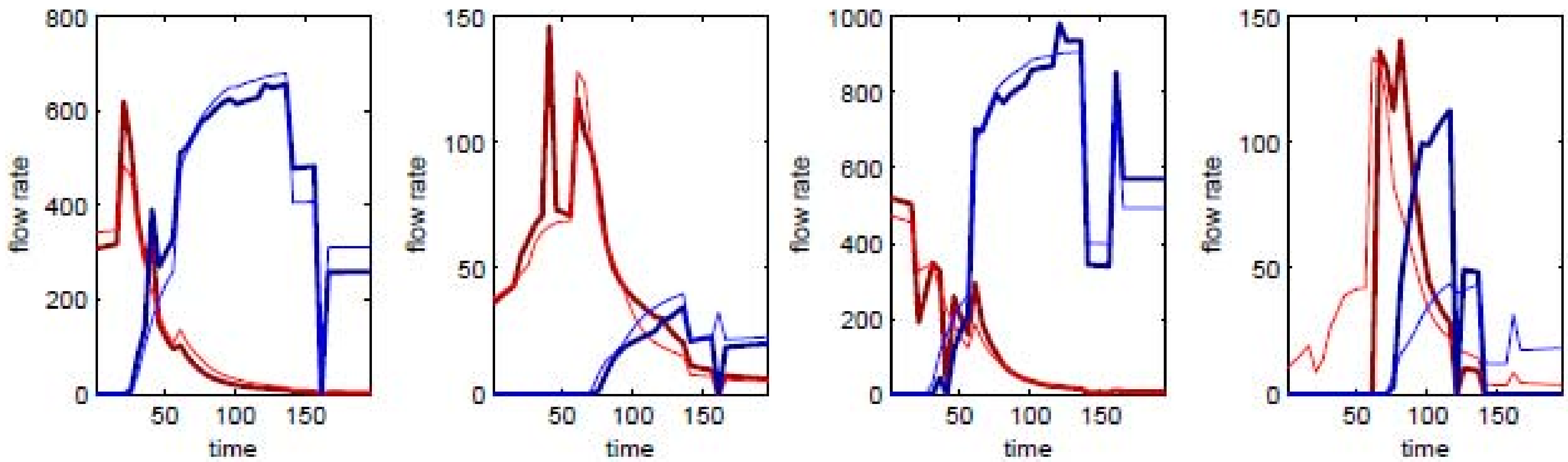
initial estimate (local point)

final estimate

Observation:

Only grid block permeabilites around well are identifiable.

Simple reservoir example



Simulated production of estimated model (thin lines) of water (blue) and oil (red) in the four producer wells

See also Vasco et al. (1997)

Simple reservoir example

- Model estimation is done in an iterative way:
  - Choosing a local identifiable parametrization
  - Estimating the parameters
  - Repeating the procedure until convergence of the cost function
- During the iterations, the quadratic cost function is reduced from 135 to 5.93.
- “Poor” model seems to be good enough for prediction of production.
- No prior info on permeability structure has been used.

Relation with controllability and observability

- Does (local) identifiability relate to properties of observability and controllability?

$$\mathbf{h}_k(\boldsymbol{\theta}) = \mathbf{C}(\boldsymbol{\theta})\mathbf{x}_k, \qquad \mathbf{x}_{k+1} = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}_k + \mathbf{B}(\boldsymbol{\theta})\mathbf{u}_k$$

- Without loss of generality: $\mathbf{C}(\boldsymbol{\theta}) = \mathbf{C}$

$$\frac{\partial \mathbf{h}_k(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}(i)} = \mathbf{C}\frac{\partial \mathbf{x}_k}{\partial \boldsymbol{\theta}(i)},$$

$$\frac{\partial \mathbf{x}_{k+1}}{\partial \boldsymbol{\theta}(i)} = \mathbf{A}(\boldsymbol{\theta})\frac{\partial \mathbf{x}_k}{\partial \boldsymbol{\theta}(i)} + \underbrace{\frac{\partial \mathbf{A}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}(i)}\mathbf{x}_k + \frac{\partial \mathbf{B}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}(i)}\mathbf{u}_k}_{:=\tilde{\mathbf{u}}_k^{\boldsymbol{\theta}(i)}}$$

Substitution of the two equations delivers:

Relation with controllability and observability

$$\left(\frac{\partial \mathbf{h}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}}\right)^T = \underbrace{\begin{bmatrix} \mathbf{C} & & & \mathbf{0} \\ & \mathbf{C} & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} & \cdots & & \\ \mathbf{A} & \mathbf{I} & & & \\ \mathbf{A}^2 & \mathbf{A} & \mathbf{I} & & \\ \vdots & & & \ddots & \\ \mathbf{A}^{N-1} & \mathbf{A}^{N-2} & \cdots & \mathbf{A} & \mathbf{I} \end{bmatrix}}_{\tilde{\mathcal{O}} \in \mathbb{R}^{N(p \times n)}}$$

$$\times \underbrace{\begin{bmatrix} \tilde{\mathbf{u}}_0^{\boldsymbol{\theta}(1)} & \cdots & \tilde{\mathbf{u}}_0^{\boldsymbol{\theta}(i)} & \cdots & \tilde{\mathbf{u}}_1^{\boldsymbol{\theta}(q)} \\ \tilde{\mathbf{u}}_1^{\boldsymbol{\theta}(1)} & \cdots & \tilde{\mathbf{u}}_1^{\boldsymbol{\theta}(i)} & \cdots & \tilde{\mathbf{u}}_2^{\boldsymbol{\theta}(q)} \\ \vdots & & \vdots & & \vdots \\ \tilde{\mathbf{u}}_{N-1}^{\boldsymbol{\theta}(1)} & \cdots & \tilde{\mathbf{u}}_{N-1}^{\boldsymbol{\theta}(i)} & \cdots & \tilde{\mathbf{u}}_{N-1}^{\boldsymbol{\theta}(q)} \end{bmatrix}}_{\tilde{\mathbf{U}} \in \mathbb{R}^{Nn \times q}},$$

With abuse of notation:

$$\left(\frac{\partial \mathbf{h}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}}\right)^T = \tilde{\mathcal{O}} \begin{bmatrix} \frac{\partial \mathbf{A}}{\partial \boldsymbol{\theta}} & \frac{\partial \mathbf{B}}{\partial \boldsymbol{\theta}} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \tilde{\mathbf{u}} \end{bmatrix}$$

Relation with controllability and observability

$$\left(\frac{\partial \mathbf{h}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}}\right)^T = \tilde{\mathcal{O}} \begin{bmatrix} \frac{\partial \mathbf{A}}{\partial \boldsymbol{\theta}} & \frac{\partial \mathbf{B}}{\partial \boldsymbol{\theta}} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix}$$

Effect of current state and input, related to controllability

Sensitivity of system matrices with respect to parameter perturbations

Effect of state perturbation on output, related to observability

Reservoir dynamics live in low-order space

- Observation and control in the wells
  - Models will typically be poorly observable and/or poorly controllable
  - Real (local) input-output dynamics is of limited order

- Parameter estimation:
  - Physical parameters (permeabilities) determine predictive quality but one parameter per grid block leads to excessive over-parametrization (not to be validated)

Contents

- Introduction – closed-loop reservoir management
- Model estimation
- Identifiability and model structure approximation
- Simple example
- Extension to data driven approaches: two-stage procedure
- Summary

Extension: a two stage approach

Reasoning

- Optimization on the basis of nonlinear reservoir models suffers from model uncertanties
- Optimization on the basis of estimated models suffers from a lack of predictive capabilities beyond the –local- measurement interval

→ Combine the two

Extension: a two stage approach [Van Essen, 2010]

• Combine data-driven estimation of local (linear) models with tracking of long-term production targets

• Use nonlinear reservoir model (with estimated/prior chosen) grid block parameters for an (slow) ”outer loop” optimization target strategy

• Base short term operational decisions to follow the target strategy on a locally identified linear model, using simple black box estimation techniques, on the basis of deliberately perturbed input settings

Remarks

- Base short term actions on short term models
- They are (locally) more reliable and can be validated
- But are limited in their long-term predictions

- Locally identified black-box linear models, could be replaced by
  - limited complexity (identifiable, nonlinear) physical models
  - simple nonlinear black box extensions (LPV)

- Model uncertainties need to be quantified, and incorporated in the (long term) optimization strategy [Van Essen et al, SPE J, 2009]

Summary

- The development / handling of appropriate models is a key issue in closed-loop reservoir management.
- In comparison with process technology, limiting features include the nonlinear and batch-type of the process, together with highly uncertain process knowledge
- Large-scale reservoir models are not identifiable from production data, nor can they be validated
- A proper balance between data-induced and priors-induced modelling should be achieved when estimating models, focussing at the control-relevant dynamics
- Model uncertainty should be specified and incorporated in the optimization at all levels; a separation of time scales is one of the options to support this

With thanks to our sponsors:

Consortia: VALUE (Shell, TU Delft, MIT), ISAPP (Shell, TNO, TU Delft),

Recovery Factory (Shell, TUD)