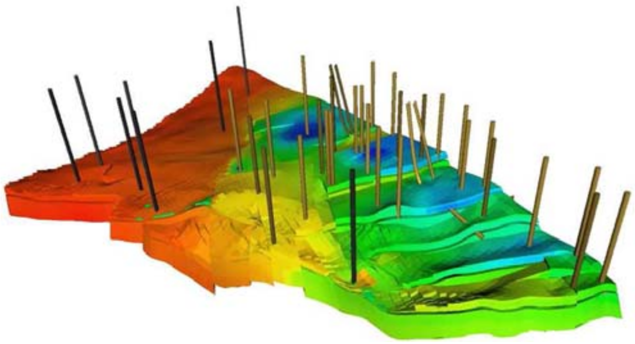


Identification of parameters in large-scale physical model structures

Motivated by an example from reservoir engineering

Paul M.J. Van den Hof

co-authors: Jorn Van Doren, Jan Dirk Jansen, Sippe Douma Okko Bosgra

Tsinghua University, Beijing, China, 19 October 2009

Delft University of Technology

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

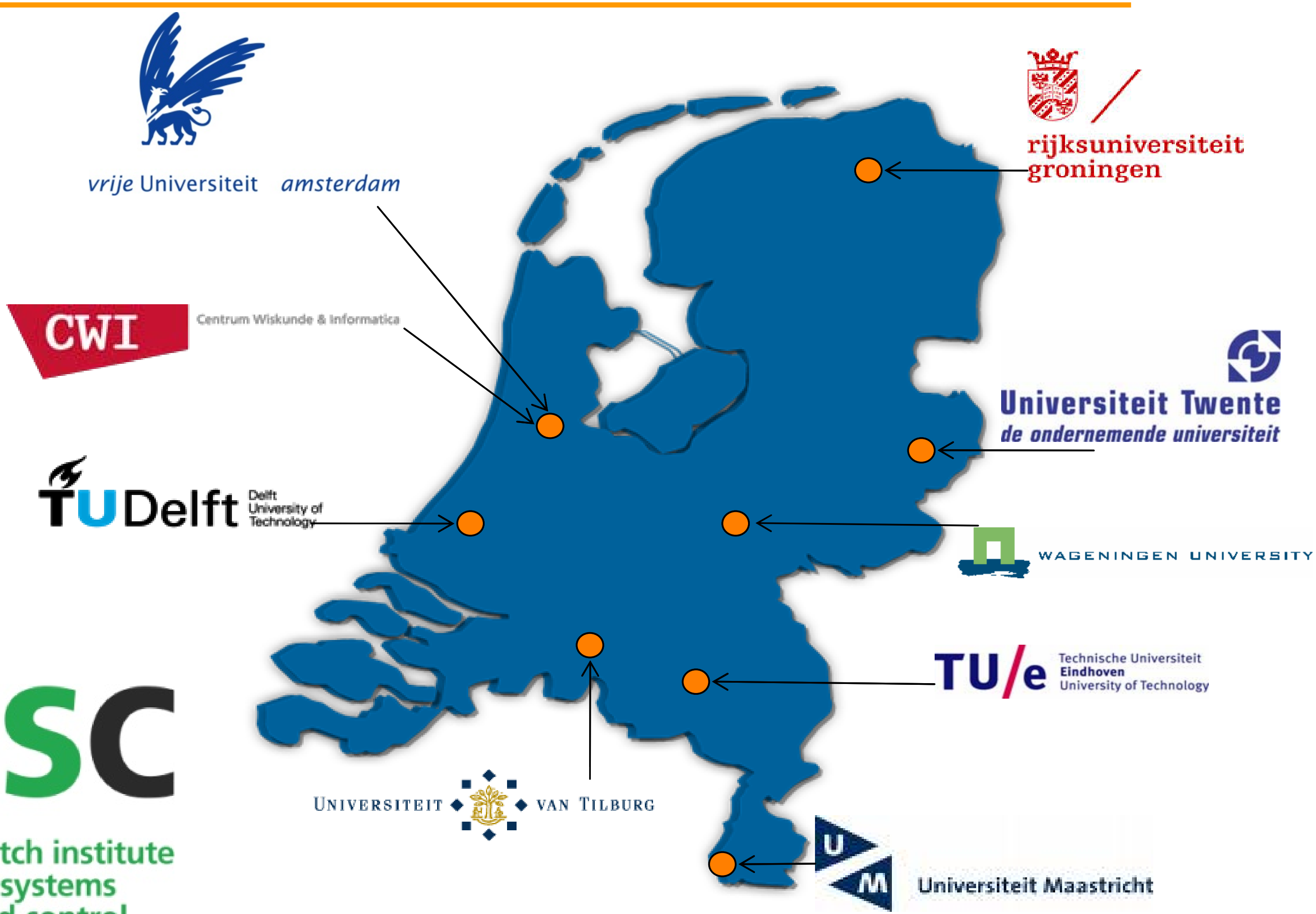
Founded in 1842 as an engineering school

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

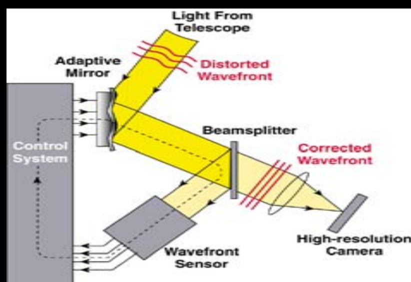
Delft

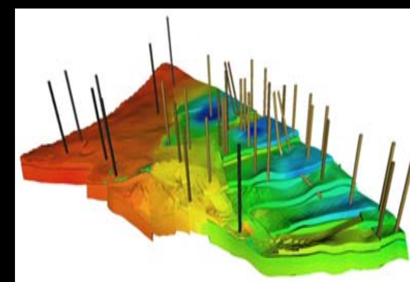
Dutch Universities in Control



disc
dutch institute of systems and control

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:

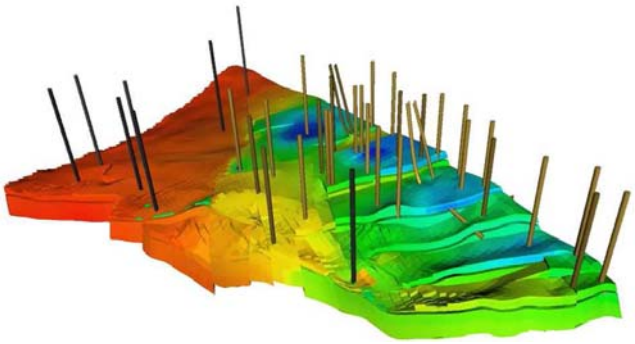
- Automotive systems
- Microfactory
- AFM nano-positioning
- Smart optics systems
- Electron microscopes
- Robotics

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



Identification of parameters in large-scale physical model structures

Motivated by an example from reservoir engineering

Paul M.J. Van den Hof

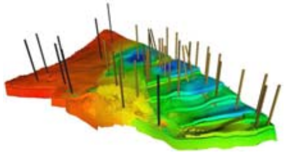
co-authors: Jorn Van Doren, Jan Dirk Jansen, Sippe Douma Okko Bosgra

Tsinghua University, Beijing, China, 19 October 2009

Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- Identifiability
- Model structure approximation
- Parameter scaling
- Particular connections:
  - Observability and controllability
  - Bayesian estimation
  - Structural identifiability
- Simple reservoir example
- Discussion

Motivation:
physical structures versus black box



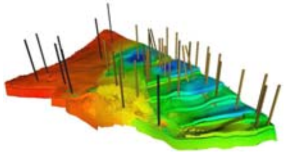
- Black box model structures (transfer function, state-space etc.) are well suited for identification of linear systems.

- For nonlinear systems there are multiple approaches possible, and a choice for an efficient model structure will be highly case-dependent (neural nets, Wiener/Hammerstein, polynomials, Volterra kernels....)

- Typical issue in model structure selection: strive towards as few parameters as possible/necessary.

Rough definition:

Model structure $\mathcal{M} : \boldsymbol{\theta} \rightarrow i/o\ map$



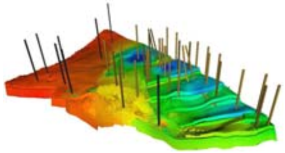
Motivation: physical structures versus black box

- Which mechanisms can induce a limited number of parameters?
  - *Sheer luck* : guess a model structure that needs only a few parameters to approximate the system well.
  - *Prior information* : use information on the structure / dynamics of the underlying system

- A model structure / parametrization parametrizes a set of models that should reflect the prior uncertainty that we have about our system to be identified.

If that prior info is: “*any nonlinear system*”
the identification task becomes huge!

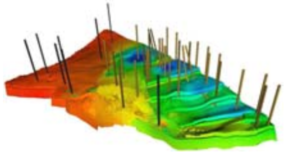
Motivation:
physical structures versus black box



- Physical model structures can help us to
    - Come up with model structures with *limited number of parameters*.
    - *Extrapolate* the model dynamics to (nonlinear) regimes that are not necessarily captured in the data.

      Rather than linearizing a model around an operating point, estimate the parameters in a (reliable) nonlinear model with a predefined (physics-motivated) structure.

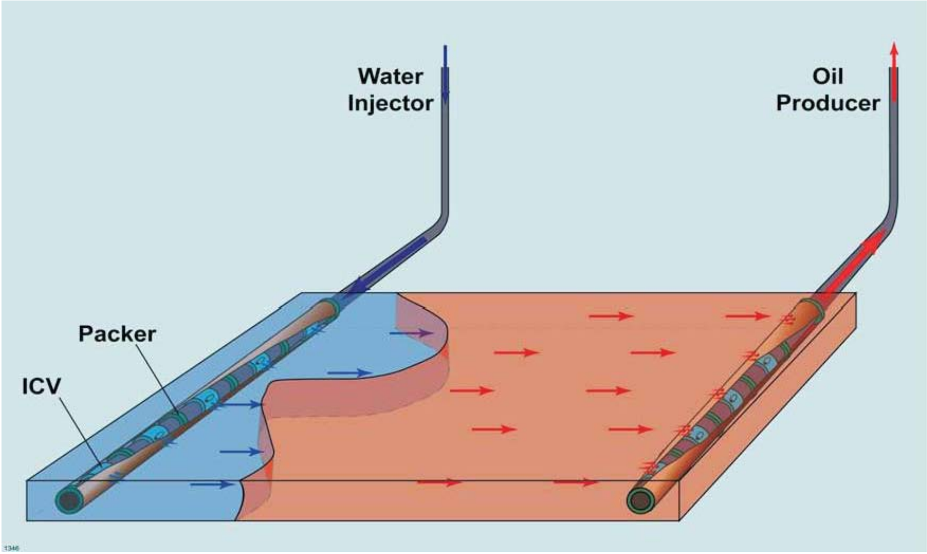
- Physical model structures raise the problem of identifiability:
  can all unknown parameters be estimated uniquely?



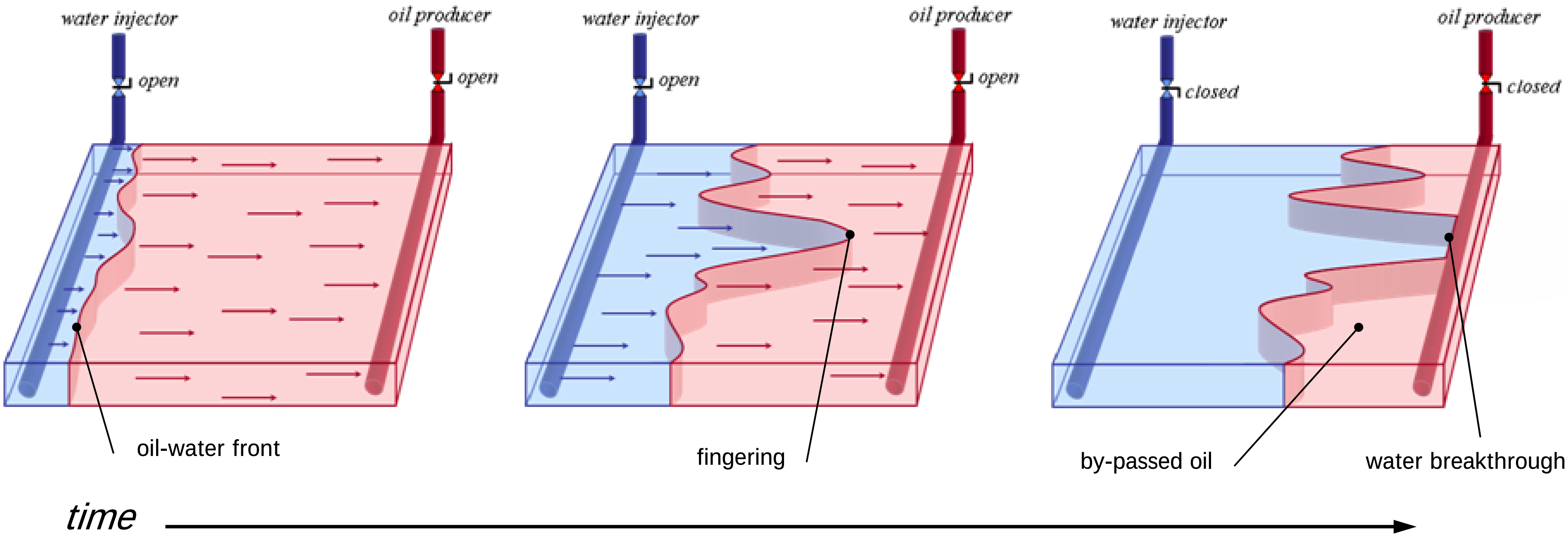
Motivation:
an example from oil reservoir engineering

Water flooding

- Involves the injection of water through the use of injection wells
- Goal is to displace oil by water
- Production is terminated when (too much) water is being produced



Motivation:
an example from oil reservoir engineering



Oil production from a reservoir can last for periods of 15-25 years.

Motivation:
an example from oil reservoir engineering

Mass balance:

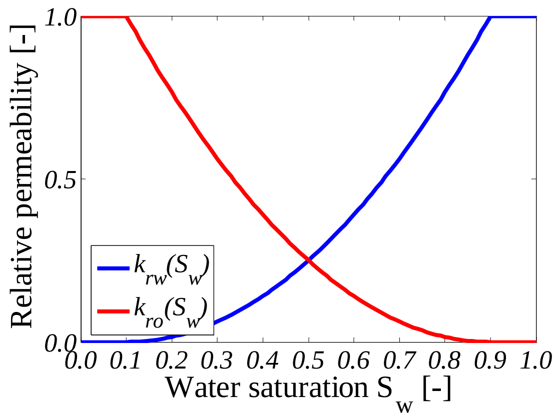
$$\nabla(\rho_i u_i) + \frac{\partial}{\partial t}(\phi \rho_i S_i) = 0 \qquad i = \{o, w\}$$

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

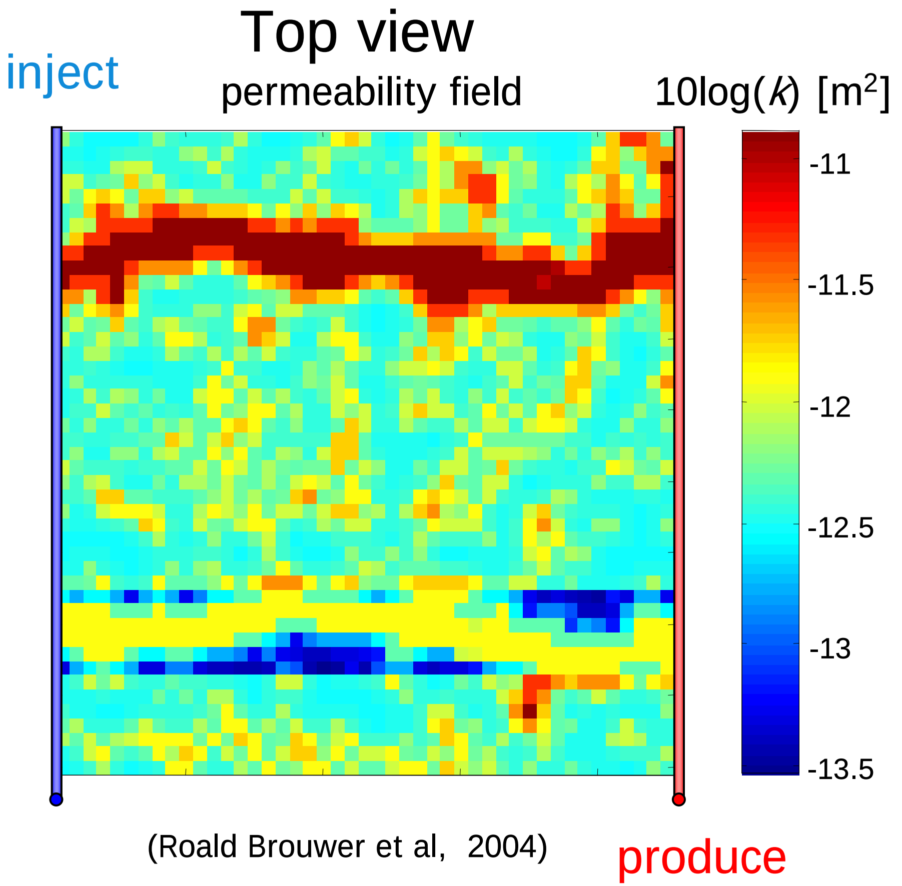
nonlinearity



After discretization in space (grid blocks):
a nonlinear ODE results, with
separate permeability parameters in each
grid block → typically $\approx 10^4 - 10^6$

Example permeability structure of a toy-reservoir:

- 2 horizontal smart wells
  - 45 inj. & prod. segments
- 2 higher permeability streaks
- Dynamics is dependent on position of water-oil front (nonlinearity)
- Typical non-linear batch process with extremely many (unknown) parameters.



(Roald Brouwer et al, 2004)

Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- **Identifiability**
- Model structure approximation
- Parameter scaling
- Particular connections:
  - Observability and controllability
  - Bayesian estimation
  - Structural identifiability
- Simple reservoir example
- Discussion

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 $\mathbf{y} := [y_1^T \cdots y_N^T]^T$

- 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]

- Structural identifiability in similar way on the basis of transfer functions (rather than output), for the linear(ized) dynamics case.

$$\{G(z, \boldsymbol{\theta}_1) = G(z, \boldsymbol{\theta}_2)\} \Rightarrow \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$$

  [Bellman and Åström (1970)]

- Global properties are generally hard to analyze

Identifiability

- 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 identification

- In Prediction Error framework, identification criterion

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

- 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

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

- If $\boldsymbol{\Sigma}_\mathbf{2} = 0$ 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

Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- Identifiability
- **Model structure approximation**
- Parameter scaling
- Particular connections:
  - Observability and controllability
  - Bayesian estimation
  - Structural identifiability
- Simple reservoir example
- Discussion

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 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 $V(\boldsymbol{\theta})$:

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

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

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

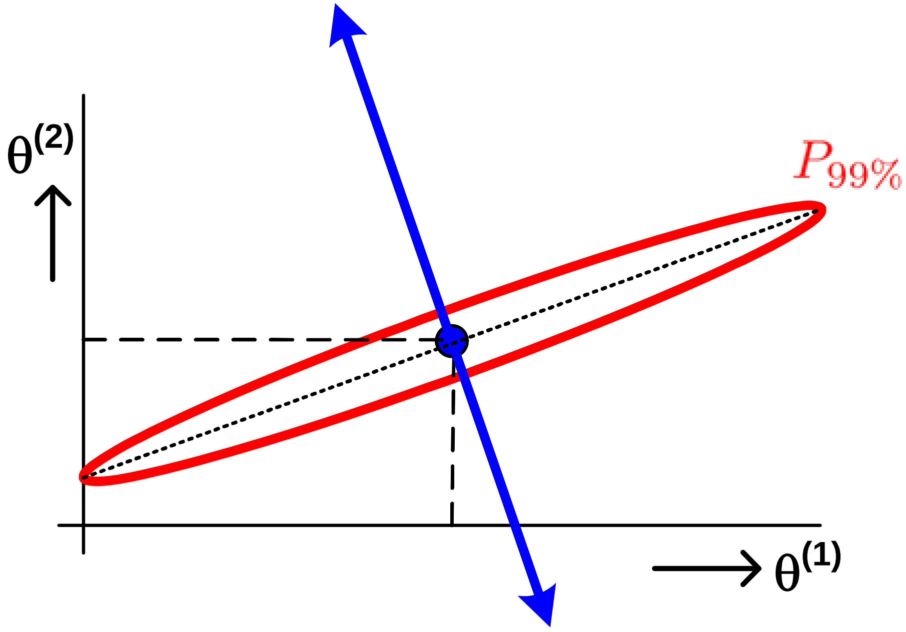
Approximating the identifiable parameter space

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

- 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]

Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- Identifiability
- Model structure approximation
- **Parameter scaling**
- Particular connections:
  - Observability and controllability
  - Bayesian estimation
  - Structural identifiability
- Simple reservoir example
- Discussion

Effect of parameter scaling/units

- In physical model structures there is a freedom of choice in parameter units (cm,km .....)
- Choice of units (scaling) should not influence the choice for approximation of the model structure!

Toy example

Second order system: $y(t) = \alpha_0 u(t-1) + \beta_0 u(t-2)$

$$\alpha_0 = 10^6; \quad \beta_0 = 10^{-6}$$

Second order FIR model:

$$\hat{y}(t,\theta) = \alpha u(t-1) + \beta u(t-2), \quad \theta := [\alpha \quad \beta]^T.$$

$$\psi(t,\theta_0) := \frac{\partial \hat{y}(t,\theta)}{\partial \theta} = \begin{bmatrix} u(t-1) \\ u(t-2) \end{bmatrix}$$

Fisher information matrix: $J = N \begin{bmatrix} R_u(0) & R_u(1) \\ R_u(1) & R_u(0) \end{bmatrix} = N \cdot I$

unit variance white noise input; $\mathbf{P}_v = 1$

→ No indications for reducing model structure

Toy example (cnt'd)

Second order system: $y(t) = \alpha_0 u(t-1) + \beta_0 u(t-2)$

$$\alpha_0 = 10^6; \quad \beta_0 = 10^{-6}$$

However, if we reparametrize:

$$\hat{y}(t,\theta) = \alpha u(t-1) + 10^{-6}\gamma u(t-2), \quad \theta := [\alpha \ \ \gamma]^T.$$

$$\psi(t,\theta_0) := \frac{\partial \hat{y}(t,\theta)}{\partial \theta} = \begin{bmatrix} u(t-1) \\ 10^{-6}u(t-2) \end{bmatrix}$$

Fisher information matrix: $J = N \begin{bmatrix} 1 & 0 \\ 0 & 10^{-12} \end{bmatrix}$

→ Strong indication for removing the γ parameter

Effect of parameter scaling/units

- Scaling-dependence of the results is very undesirable!

What are the observed mechanisms here?

- The yes/no question on identiability:

$$\mathbf{\Sigma_2} = 0, \quad \mathbf{\Sigma_2} \neq 0$$

is not influenced by a scaling of parameter values:

$$\hat{\theta} = \mathbf{\Gamma}\hat{\theta}_1, \quad \mathbf{\Gamma} = \mathrm{diag}(s_1, \cdots s_n)$$

- However for $\mathbf{\Sigma_2} \neq 0$, scaling will influence the numerical values of $\mathbf{\Sigma_1}, \mathbf{\Sigma_2}$ and therefore also the choice of the identifiable parameter space

Effect of parameter scaling/units

- The Fisher matrix considers an absolute parameter variance measure

  While a parameter variance of 1 has different consequences for a nominal parameter of 10^6 or 10^{-6}

- The remedy here is: defining a measure for the relative variance

Effect of parameter scaling/units

- Possible remedy: use relative parameter variance rather than absolute variance as a measure for model structure approximation

$$\mathrm{cov}(\boldsymbol{\Gamma}_{\hat{\theta}}^{-1}\hat{\boldsymbol{\theta}}), \quad \text{e.g. } \boldsymbol{\Gamma}_{\hat{\theta}} = \mathrm{diag}\left(\; |\hat{\boldsymbol{\theta}}_1| \quad \ldots \quad |\hat{\boldsymbol{\theta}}_q| \; \right)$$

- Motivates analysis of scaled Hessian $\boldsymbol{\Gamma}_{\hat{\theta}} \left.\frac{\partial^2 V(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right|_{\hat{\boldsymbol{\theta}}} \boldsymbol{\Gamma}_{\hat{\theta}}$

- Essential information in

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

- Model structure approximated: parameters identifiable and physically interpretable!

Consequence for toy example

Second order system: $y(t) = \alpha_0 u(t-1) + \beta_0 u(t-2)$

$$\alpha_0 = 10^6; \quad \beta_0 = 10^{-6}$$

Second order FIR model:

$$\hat{y}(t,\theta) = \alpha u(t-1) + \beta u(t-2), \quad \theta := [\alpha \;\; \beta]^T.$$

Scaled Fisher information matrix:

$$\begin{aligned} \tilde{J} &= N \begin{bmatrix} \alpha_0 & 0 \\ 0 & \beta_0 \end{bmatrix} \begin{bmatrix} R_u(0) & R_u(1) \\ R_u(1) & R_u(0) \end{bmatrix} \begin{bmatrix} \alpha_0 & 0 \\ 0 & \beta_0 \end{bmatrix} \\ &= N \begin{bmatrix} 10^{12} & 0 \\ 0 & 10^{-12} \end{bmatrix} \end{aligned}$$

→ Indication for reducing model structure by one parameter

and this result is invariant for scaling of any of the parameters.

Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- Identifiability
- Model structure approximation
- Parameter scaling
- Particular connections:
  - **Observability and controllability**
  - Bayesian estimation
  - Structural identifiability
- Simple reservoir example
- Discussion

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

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

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

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

Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- Identifiability
- Model structure approximation
- Parameter scaling
- Particular connections:
  - Observability and controllability
  - **Bayesian estimation**
  - Structural identifiability
- Simple reservoir example
- Discussion

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 $\mathbf{P}_{\boldsymbol{\theta}_p}$).

- Model output approximated with first-order Taylor expansion. 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

- 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)

Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- Identifiability
- Model structure approximation
- Parameter scaling
- Particular connections:
  - Observability and controllability
  - Bayesian estimation
  - **Structural identifiability**
- Simple reservoir example
- Discussion

Structural identifiability

- The analysis presented so far, can similarly be given for local structural identifiability (in stead of local identifiability).

- Rather than the system output (which is input dependent) we then consider the transfer function:

$$G(z,\boldsymbol{\theta}) = \sum_{k=1}^{\infty} \mathbf{M}(k,\boldsymbol{\theta}) z^{-k}$$

- Structurally identifiable in $\boldsymbol{\theta}_m$ if $\dfrac{\partial \mathbf{S}_N^T(\theta)}{\partial \boldsymbol{\theta}}$ full rank, where

$$\mathbf{S}_N(\boldsymbol{\theta}) := [\ \mathbf{M}(1,\boldsymbol{\theta}) \quad \dots \quad \mathbf{M}(N,\boldsymbol{\theta})\ ]^T$$

- The matrix to be quantitatively analyzed becomes: $\boldsymbol{\Gamma}_{\theta_m} \dfrac{\partial \mathbf{S}_N^T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \boldsymbol{\Gamma}_S$

Structural identifiability

- For linear(ized) models, a closed-form expression exists for

$$\frac{\partial \mathbf{S}_N^T(\theta)}{\partial \boldsymbol{\theta}}$$

on the basis of:

$$\mathbf{A}(\boldsymbol{\theta}), \mathbf{B}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}), \frac{\partial \mathbf{A}}{\partial \boldsymbol{\theta}}, \frac{\partial \mathbf{B}}{\partial \boldsymbol{\theta}}, \frac{\partial \mathbf{C}}{\partial \boldsymbol{\theta}}$$

[Van Doren et al., IFAC World Congress 2008]

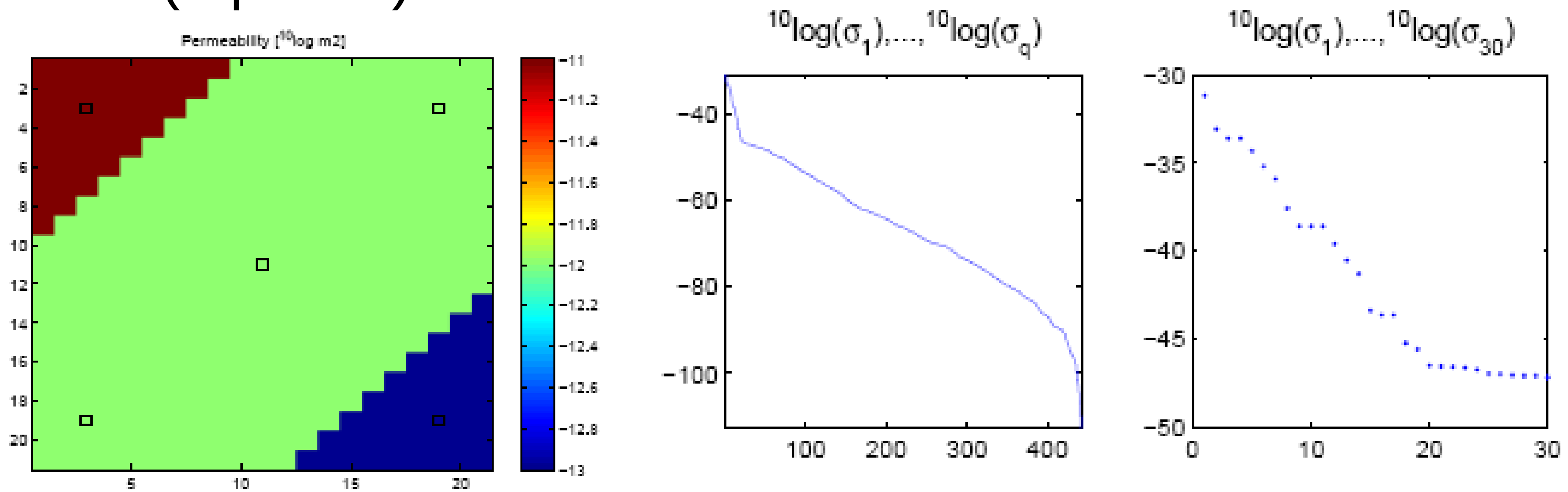
Contents

- Motivation:
  - physical structures versus black box
  - an example from reservoir engineering
- Identifiability
- Model structure approximation
- Parameter scaling
- Particular connections:
  - Observability and controllability
  - Bayesian estimation
  - Structural identifiability
- **Simple reservoir example**
- Discussion

Simple reservoir example

Single phase example

(top view)



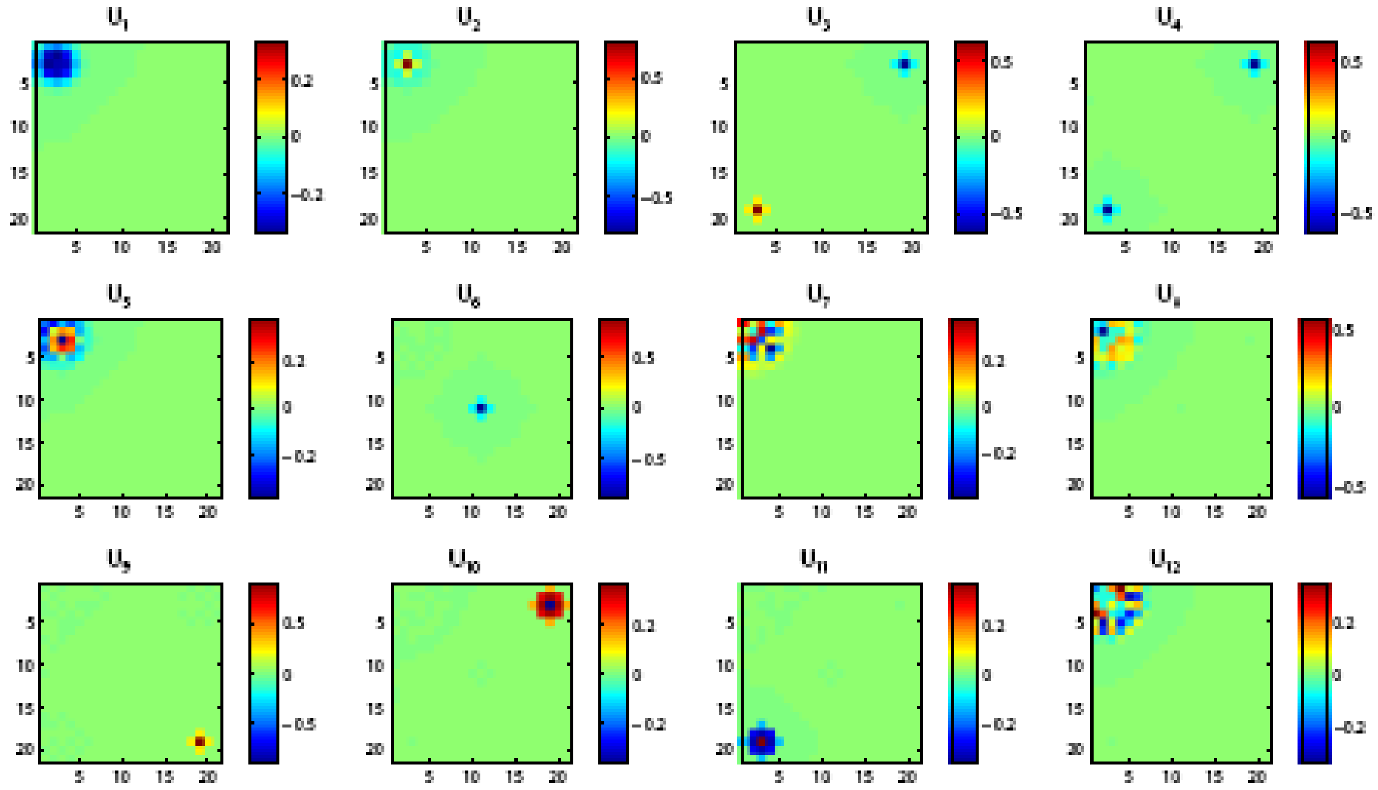
21 x 21 grid block permeabilities

5 wells; 3 permeability strokes

Delft Center for Systems and Control

Simple reservoir example

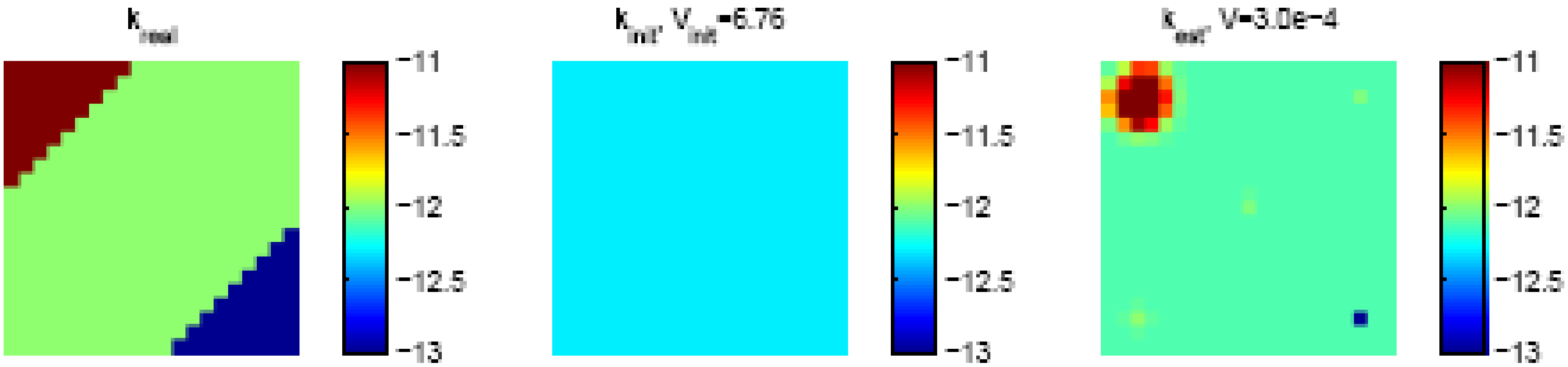
Singular vectors can be projected on the grid:



First 12 singular vectors; structural identifiability case

Simple reservoir example

Using the reduced parameter space –iteratively-
in identification:



exact field | initial estimate (local point) | final estimate

Observation:

Only grid block permeabilites around well (in high permeable area) are identifiable.

Simple reservoir example

- Same observation holds true for the case of oil/water (2-phase flow) -- (results under progress)
- Grid block properties far away from wells are poorly identifiable
- There are indications that they might not be very important for the optimal control strategy......

Discussion

- Estimating physical parameters in large-scale models is highly relevant
- Notion of identifiability quantified
- Model structure approximated to achieve identifiability, while retaining interpretation of physical parameters
- Analysis can only be done locally linearized
- Established relation with
  - Controllability and observability properties
  - Bayesian estimation
- Illustrating example (ungoing work) in reservoir engineering

Cooperative project, together with:

Jorn Van Doren, PhD student

Jan Dirk Jansen
Dept. of Geotechnology TUD and Shell Intern. Exploration and Production

Sippe Douma, Shell, and former PhD student

Okko Bosgra, DCSC, TUD

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

THE VALUE OF SMARTNESS