ORIGINAL PAPER

A system-theoretical approach to selective grid coarsening of reservoir models

S. A. Vakili-Ghahani · J. D. Jansen

Abstract From a system-theoretical point of view and for a given configuration of wells, there are only a limited number of degrees of freedom in the input–output dynamics of a reservoir system. This means that a large number of combinations of the state variables (pressure and saturation values) are not actually controllable and observable from the wells, and accordingly, they are not affecting the input–output behavior of the system. In an earlier publication, we therefore proposed a control-relevant upscaling methodology that uniformly coarsens the reservoir. Here, we present a control-relevant selective (i.e. non-uniform) coarsening (CRSC) method, in which the criterion for grid size adaptation is based on ranking the grid block contributions to the controllability and observability of the reservoir system. This multi-level CRSC method is attractive for use in iterative procedures such as computer-assisted flooding optimization for a given configuration of wells. In contrast to conventional flow-based coarsening techniques our method is independent of the specific flow rates or pressures imposed at the wells. Moreover the system-theoretical norms employed in our method provide tight upper bounds to the 'input–output energy' of the fine and coarse systems. These can be used as an a priori error-estimate of the performance of the coarse model. We applied our algorithm to two numerical examples and found that it can accurately reproduce results from the corresponding fine-scale simulations, while significantly speeding up the simulation.

Keywords Upscaling · Grid coarsening · Grid refinement · Controllability · Observability · Control-relevant · Hankel singular values · System theory

Nomenclature

A	System matrix
B	Input matrix
C	Output matrix
D	Direct throughput matrix
e	Relative error
E	Energy
I	Identity matrix
J	Well index matrix, $m^3/(Pa\ s)$
k	Counter (discrete time)
K	Absolute permeability, m^2
l	Number of snapshots
ℓ	Cut-off value for singular values
L	Cholesky factor
n	Number of states; number of grid blocks in fine-scale model
p	Pressure vector, Pa
q	Well flow rate vector, m^3/s
t	Time, s
T	Transformation matrix; transmissibility matrix, $m^3/(Pa\ s)$
u	Input vector; singular vector

S. A. Vakili-Ghahani · J. D. Jansen (✉)
Delft University of Technology, Stevinweg 1, 2628 CN, Delft, The Netherlands
e-mail: j.d.jansen@tudelft.nl

S. A. Vakili-Ghahani
e-mail: s.a.vakilighahani@tudelft.nl

U	Matrix of left singular vectors
V	Matrix of right singular vectors; accumulation matrix
W	Gramian
x	State vector
X	Snapshot matrix for the forward model
y	Output vector
Y	Snapshot matrix for the adjoint model
z	Adjoint state vector
α	Threshold value
ε	Machine epsilon
σ	Singular value
Σ	Diagonal matrix containing singular values

Subscripts

a	Adjoint
c	Continuous
cum	Cumulative
C	Controllability
E	Energy
HSV	Hankel singular values
k	Counter (discrete time)
l	Counter (grid blocks)
O	Observability
p	Pressure
q	Flow rate
w	Well

Superscripts

| r | Order of low-rank approximated Cholesky factors |
| T | Transpose |

1 Introduction

1.1 Conventional upscaling

At present, the computational limits for reservoir flow simulation restrict the model order typically to 10^4 to 10^6, for single-core implementations. Here, the model order is defined as the number of time-dependent variables (i.e. state variables such as grid block pressures, saturations or component accumulations) which is typically equal to the number of active grid blocks times the number of components (i.e. hydrocarbon components and water) in the simulation. The number of time-independent model parameters is usually of the same order of magnitude because they are also proportional to the number of grid blocks (e.g. grid block permeabilities and porosities). However, geological subsurface models often represent the subsurface heterogeneity with 10^6 to 10^9 parameters ('voxels') and some form of

upscaling is therefore required to transfer the relevant features of a geological model to a flow simulation model. The uncertainty of the geological parameters is increasingly taken into account by simulating an ensemble of model realizations, which significantly increases the computational demands, especially when it is also required to perform repeated simulations for iterative procedures such as computer-assisted flooding optimization or history matching. In particular, we consider the application of reservoir simulation for 'closed-loop reservoir management', which involves the use of simulation models during the producing life of a reservoir for near-continuous flooding optimization based on frequently updated, 'evergreen', reservoir models [8, 20, 21, 27, 32]. Even although the rapid increase of cluster computing facilitates such operational use of reservoir simulation, reducing the number of grid blocks, and thus the model order, through upscaling remains a computational and practical necessity. There are different grid-based upscaling techniques varying from simple averaging methods on uniform Cartesian cells to sophisticated flow-based techniques on adaptive and unstructured grids. Extensive reviews of the different methods were written by, e.g., [10, 11, 30, 41].

1.2 Control-relevant upscaling

In most conventional upscaling techniques, the coarse-scale parameters are calculated based on the fine-scale parameters and some local flow calculations, but not any other system property. However, we argue that, from a system-theoretical point of view, there are only a limited number of degrees of freedom in the input–output dynamics of a reservoir system [43]. This means that, for a given configuration of wells, a large number of combinations of the state variables (pressure and saturation values) are not actually controllable and observable from the wells, and accordingly, they are not affecting the input–output behavior of the system. Therefore, the complexity level of a model should be adjusted to the amount of available information and the extent of control that is possible in the reservoir system. In an earlier publication, we therefore proposed a control-relevant upscaling (CRU) method that uniformly coarsens the reservoir model based on the relevant level of information and control (see [37]). Note that we use the terms upscaling and coarsening interchangeably.

The CRU method is a global method, in the sense that it relies on the system properties of the entire reservoir. It does not, however, require any forward simulation, neither of the full nor of the upscaled model. It also does not depend on a particular control

strategy, but instead uses the dynamical system equations directly. Its dependency on well locations, however, implies that it should be (partially) repeated when those locations are changed. Moreover, the formulation as presented in Vakili-Ghahani and Jansen [37] is restricted to fine-scale models with a maximum of around 10^5 grid blocks because of current limits on the computation of the underlying system norms. In Vakili-Ghahani et al. [35] we proposed to overcome this computational limit by combining CRU with model-order reduction techniques. In Vakili-Ghahani and Jansen [36], we followed a different route and proposed a two-level control-relevant selective coarsening (CRSC) algorithm. Here we will treat the latter approach more in-depth, extend it to multiple levels of coarsening, include alternative computational methods, perform an operations count analysis, and add a second numerical example.

The multi-level CRSC algorithm selectively refines the grid to various degrees in the most controllable and observable parts of the reservoir. Alternatively, it starts from a uniformly coarsened grid and then selectively refines it. In addition to a computational advantage, the selective coarsening/refinement is also expected to be more accurate, in the sense that a selectively upscaled model is expected to introduce smaller upscaling errors than a uniformly upscaled model with the same number of grid blocks. The main contribution of this paper is to present a control-relevant criterion for grid adaptation, and use it in a selective (non-uniform) coarsening/refinement approach that is applicable to large fine-scale models. In the next sections we will first provide some background information on system-theoretical concepts and definitions. Thereafter we will explain the CRSC algorithm and demonstrate its performance using two numerical examples.

2 System-theoretical concepts

Some basic system-theoretical notations and their definitions are briefly discussed in this section. The system-theoretical material presented here is well-established, see, e.g. [2, 33]. For recent applications of system theory to reservoir modeling see [6, 14, 19, 24, 25, 34, 37, 38, 43].

2.1 State-space representation of the reservoir model

Consider an isothermal weakly compressible single-phase flow model with a given set of boundary and well conditions (e.g. [3]). Spatial discretization of the original partial differential equations, after some manipulation, results in

$$
\begin{bmatrix} \mathbf{V}_{11} & 0 & 0 \\ 0 & \mathbf{V}_{22} & 0 \\ 0 & 0 & \mathbf{V}_{33} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{p}}_1 \\ \dot{\mathbf{p}}_2 \\ \dot{\mathbf{p}}_3 \end{bmatrix} + \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \mathbf{T}_{13} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \mathbf{T}_{23} \\ \mathbf{T}_{31} & \mathbf{T}_{32} & \mathbf{T}_{33} + \mathbf{J}_p \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix}
$$
$$
= \begin{bmatrix} \mathbf{0} \\ \check{\mathbf{q}}_w \\ \mathbf{J}_p \check{\mathbf{p}}_w \end{bmatrix}, \tag{1}
$$

where the diagonal block matrices $\mathbf{V}_{ii}$, $i = 1, 2, 3$ are accumulation matrices with entries that depend on the grid block size, grid block porosities, and the total compressibility, and the band-diagonal block matrices $\mathbf{T}_{ij}$, $i = 1, 2, 3$, $j = 1, 2, 3$ are transmissibility matrices with entries that depend on the grid block size, grid block permeabilities, and the fluid viscosity. The elements of vector $\mathbf{p}_1$ are the pressures in those grid blocks (elements) that are not penetrated by a well. The elements of $\mathbf{p}_2$ are the pressures in the blocks where the source terms are prescribed well flow rates $\check{\mathbf{q}}_w$, and those of $\mathbf{p}_3$ are the pressures in the blocks where the source terms are obtained through prescription of the bottom-hole pressures with the aid of a well model,

$$
\bar{\mathbf{q}}_w = \mathbf{J}_p \left(\check{\mathbf{p}}_w - \mathbf{p}_3 \right). \tag{2}
$$

Here $\mathbf{J}_p$ is a diagonal matrix of well indices, the elements of $\check{\mathbf{p}}_w$ are the prescribed pressures, and the elements of $\bar{\mathbf{q}}_w$ are the resulting well flow rates. To compute the bottom-hole pressures $\bar{\mathbf{p}}_w$ in the wells where the flow rates have been prescribed, we need an additional diagonal matrix $\mathbf{J}_q$ of well indices such that

$$
\check{\mathbf{q}}_w = \mathbf{J}_q \left(\bar{\mathbf{p}}_w - \mathbf{p}_2 \right). \tag{3}
$$

Equations 2 and 3 can be combined to give

$$
\begin{bmatrix} 0 \\ \bar{\mathbf{p}}_w \\ \bar{\mathbf{q}}_w \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{J}_p \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{J}_q^{-1} & 0 \\ 0 & 0 & \mathbf{J}_p \end{bmatrix} \begin{bmatrix} 0 \\ \check{\mathbf{q}}_w \\ \check{\mathbf{p}}_w \end{bmatrix}. \tag{4}
$$

If we define the vectors

$$
\mathbf{x} = \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix}, \tag{5}
$$

$$
\mathbf{u} = \begin{bmatrix} \check{\mathbf{q}}_w \\ \check{\mathbf{p}}_w \end{bmatrix}, \tag{6}
$$

$$
\mathbf{y} = \begin{bmatrix} \bar{\mathbf{p}}_w \\ \bar{\mathbf{q}}_w \end{bmatrix}, \tag{7}
$$

Equations 1 and 4 can be rewritten, respectively, as

$$\dot{\mathbf{x}} = \mathbf{A}_c \mathbf{x} + \mathbf{B}_c \mathbf{u}, \qquad (8)$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}, \qquad (9)$$

where

$$\mathbf{A}_c = - \begin{bmatrix} \mathbf{V}_{11}^{-1}\mathbf{T}_{11} & \mathbf{V}_{11}^{-1}\mathbf{T}_{12} & \mathbf{V}_{11}^{-1}\mathbf{T}_{13} \\ \mathbf{V}_{22}^{-1}\mathbf{T}_{21} & \mathbf{V}_{22}^{-1}\mathbf{T}_{22} & \mathbf{V}_{22}^{-1}\mathbf{T}_{23} \\ \mathbf{V}_{33}^{-1}\mathbf{T}_{31} & \mathbf{V}_{33}^{-1}\mathbf{T}_{32} & \mathbf{V}_{33}^{-1}\left(\mathbf{T}_{33} + \mathbf{J}_p\right) \end{bmatrix}, \qquad (10)$$

$$\mathbf{B}_c = \begin{bmatrix} 0 & 0 \\ \mathbf{V}_{22}^{-1} & 0 \\ 0 & \mathbf{V}_{33}^{-1}\mathbf{J}_p \end{bmatrix}, \qquad (11)$$

$$\mathbf{C} = \begin{bmatrix} 0 & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{J}_p \end{bmatrix}, \qquad (12)$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{J}_q^{-1} & 0 \\ 0 & \mathbf{J}_p \end{bmatrix}. \qquad (13)$$

Equations 8 and 9 give the standard continuous-time 'state-space representation' of a LTI system, as used in systems and control theory. The matrices $\mathbf{A}_c \in \mathbf{R}^{n \times n}$, $\mathbf{B}_c \in \mathbf{R}^{n \times m}$, $\mathbf{C} \in \mathbf{R}^{p \times n}$, and $\mathbf{D} \in \mathbf{R}^{p \times m}$ are respectively referred to as the 'system matrix', because it contains the properties of the system, the 'input matrix', because it maps the inputs to the states, the 'output matrix', because it maps the states to the outputs, and the 'direct-throughput matrix'. Here, the subscript c refers to the continuous-time form. Moreover, for a system with m inputs, p outputs and n state variables, the input vector $\mathbf{u}(t) \in \mathbf{R}^m$, where $t \in \mathbf{R}$ represents time, and the output vector $\mathbf{y}(t) \in \mathbf{R}^p$ include the flow rates and bottom-hole pressures in those wells where they can actually be controlled and/or observed. In a reservoir simulator we have access to all wells, such that m and p are both equal to the number of wells, while $\mathbf{B} = \mathbf{C}^T$. In reality control/and or observation of individual wells may not be possible, e.g. because oil is produced via a subsea manifold without individual well controls, or because injection water is distributed over a group of wells without individual rate allocation. In that case generally m $m \neq p$ and $\mathbf{B} \neq \mathbf{C}^T$. The state vector $\mathbf{x}(t) \in \mathbf{R}_+^n$ represents the state variables, i.e., for single-phase flow, the pressure values in all grid blocks. The order (or dimension) of the system is equal to n. Using an implicit time discretization with a fixed time step Δt,

Eqs. 8 and 9 can formally be rewritten in discrete-time form as

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \qquad (14)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k, \qquad (15)$$

where

$$\mathbf{A} = \left(\mathbf{I} - \Delta t \mathbf{A}_c\right)^{-1}, \qquad (16)$$

$$\mathbf{B} = \left(\mathbf{I} - \Delta t \mathbf{A}_c\right)^{-1} \Delta t \mathbf{B}_c. \qquad (17)$$

We note that the notation $\mathbf{x}_k$ indicates $\mathbf{x}(t_k)$, i.e. the value of $\mathbf{x}$ at continuous time $t = t_k$, where the subscript k is the time step counter or discrete time. Moreover, in a practical computational scheme the inverse matrices will not actually be computed, and a computationally more efficient approach will be followed. Equations 14 and 15, also known as the 'system' and 'output' equations respectively, are the standard discrete-time linear time-invariant state-space equations as used in the systems and control literature.

2.2 Balanced realization of a linear system

A stable linear time-invariant system represented by Eqs. 14 and 15 is state-controllable if we can bring it from any initial state to any final state by choosing a proper input. It is state-observable if we can reconstruct the initial state from the knowledge of the outputs. Controllability and observability can be quantified with the aid of the Gram matrices, Gramians in short, which are square and symmetric matrices, defined as

$$\mathbf{W}_C = \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{B}\mathbf{B}^T \left(\mathbf{A}^T\right)^k, \qquad (18)$$

and

$$\mathbf{W}_O = \sum_{k=0}^{\infty} \left(\mathbf{A}^T\right)^k \mathbf{C}^T \mathbf{C} \mathbf{A}^k. \qquad (19)$$

The necessary and sufficient condition for the state controllability and observability is that the respective Gramians are positive definite. Instead of having to compute the infinite sum of Eqs. 18 and 19, the Gramians can be found as solutions of Lyapunov equations

$$\mathbf{A}\mathbf{W}_C\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = \mathbf{W}_C, \qquad (20)$$

and

$$\mathbf{A}^T\mathbf{W}_O\mathbf{A} + \mathbf{C}^T\mathbf{C} = \mathbf{W}_O. \qquad (21)$$

Note that in the case that $\mathbf{B} = \mathbf{C}^T$ only one of these equations needs to be solved and $\mathbf{W}_C = \mathbf{W}_O^T$. Furthermore, to assess the degree of the system's controllability the control energy is defined as the minimum energy that is required to bring the system from the zero state to state $\mathbf{x}$, and it is given by $E_C(\mathbf{x}) = \mathbf{x}^T \mathbf{W}_C^{-1} \mathbf{x}$. Similarly, the maximum observation energy, when the system is in state $\mathbf{x}$, is defined as $E_O(\mathbf{x}) = \mathbf{x}^T \mathbf{W}_O \mathbf{x}$. The Gramians can therefore give a measure of the degree of controllability and observability of each individual state, i.e., the degree to which a state is influenced by the inputs and the effect that changes in this state have on the outputs. Note that the term 'energy' is used loosely here, motivated by the fact that energy can often be written as a quadratic form (e.g. potential energy as a function of squared pressure). A balanced realization combines the controllability and observability properties of a system by finding a transformation matrix $\mathbf{T}$ that gives an equivalent model with equal and diagonal controllability and observability Gramians, i.e. $\bar{\mathbf{W}}_C = \bar{\mathbf{W}}_O$.

As an aside, we note that the model-order reduction technique known as proper orthogonal decomposition (POD) is based on an empirical controllability Gramian obtained by collecting state snapshots from a model simulation. The balanced version of the POD method, known as BPOD, also includes an empirical observability Gramian obtained by collecting adjoint state snapshots. The diagonal entries of the balanced Gramians, known as Hankel singular values, give a ranked measure of the energy of the individual transformed states (which are linear combinations of the original states), or their contribution to the input–output behavior. Therefore, Hankel singular values can be used to identify those linear combinations of the states that represent the most important input–output characteristics of the system.

3 Control-relevant selective coarsening

The general idea behind selective (non-uniform) coarsening is to only coarsen those parts of the spatial domain that are the least important for the flow simulation according to some predefined criterion. In our CRCS application we use a control-relevant criterion based on the quantification of the controllability and observability subspaces over the spatial domain. Other selective gridding methods have been proposed, sometimes with an adaptive, i.e. time-dependent, strategy for selective refinement and coarsening based on different criteria. In particular, adaptive gridding is often applied

to maintain a fine grid in areas of high permeability or in areas where high saturation or concentration gradients occur; see, e.g., [5, 13] and references therein.

3.1 Spatial quantification of observability and controllability

As discussed in Section 2, we can quantify the controllability and the observability of a linear reservoir system by computing the Gramians $\mathbf{W}_C$ and $\mathbf{W}_O$. The square roots of the eigen values of $\mathbf{W}_C \mathbf{W}_O$ are then equal to the Hankel singular values (HSVs) of the system. The HSVs can give a measure of the combined controllability and observability of the balanced states of the system. In [35, 37] we used this analysis to develop a uniform CRU method that indirectly focuses on the most controllable and observable states of the system. There we found the coarse grid block parameters by minimizing the difference between the fine and the coarse reservoir models in terms of a system norm that characterizes the input–output behavior for a given configuration of wells. Here, in addition to the HSVs, we also consider the singular vectors. These are obtained by a singular value decomposition of

$$\mathbf{W}_C \mathbf{W}_O = \mathbf{U} \mathbf{\Sigma}^2 \mathbf{V}^T, \tag{22}$$

where $\mathbf{\Sigma}$ contains the HSVs and $\mathbf{U}$ contains the corresponding singular vectors as columns. From the relations between the Gramians and the balanced Gramians, we obtain

$$\mathbf{W}_C \mathbf{W}_O = \mathbf{T}^{-1} \bar{\mathbf{W}}_C \bar{\mathbf{W}}_O \mathbf{T} = \mathbf{T}^{-1} \mathbf{\Sigma}^2 \mathbf{T}. \tag{23}$$

Consequently, for a balanced realization of the system, the singular vectors are equivalent to the columns of $\mathbf{T}^{-1}$, where $\mathbf{T}$ is the balancing transformation matrix which results in equal and diagonal Gramians. In most reservoir models, the magnitude of HSVs are rapidly decreasing, indicating that a large number of the states are weakly controllable/observable. Hence, they weakly influence the input–output behavior of the system [34, 39, 43]. Considering this effect, we can rewrite Eq. 22 as

$$\mathbf{W}_C \mathbf{W}_O = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_1^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2^2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}, \tag{24}$$

where $\mathbf{\Sigma}_1 = diag(\sigma_1, \sigma_2, \cdots, \sigma_\ell)$ contains the first ℓ HSVs of the system (corresponding to the most controllable/observable states) which are significantly larger than the HSVs in $\mathbf{\Sigma}_2 = diag(\sigma_{\ell+1}, \sigma_{\ell+2}, \cdots, \sigma_n)$. Note that each column of $\mathbf{U}$ has n entries that are related to n states, which are in turn connected to n grid blocks. Moreover, the first ℓ columns of $\mathbf{U}$, corresponding to

the first ℓHSVs, are representing the most controllable/observable subspace of the state-space and, consequently, are related to those combination of the states that are most controllable/observable. Therefore, the singular vectors related to the first few HSVs of the system can be used to quantify the controllability and observability spatially. We select the number of HSVs that should be retained in the computations, based on the definition of the total energy transferred through the system. This energy is defined as the square root of the sum of the squared HSVs and is also referred to as the Hilbert–Schmidt–Hankel (HSH) norm [17]. However, in most cases, we only calculate ℓHSVs because the order of magnitude of the rest is below machine precision. Therefore an upper bound to the total 'input–output energy' E_{HSV} can be defined as

$$E_{\mathrm{HSV}} = \sqrt{\sum_{i=1}^{\ell} \sigma_i^2 + (n - \ell)\,\varepsilon} \qquad (25)$$

where ε represents the machine epsilon. Usually, the second term under the square root sign in Expression (25) will be much smaller than the first one so that the upper bound is very tight. Alternatively, we may choose a computationally more efficient implementation in which only a limited number ℓ of Hankel singular values is computed, i.e. not all that are above machine precision. In that case, we need to replace ε in Eq. 25 with $\sigma_{\ell+1}$ which results in a somewhat looser upper bound.

As an illustration of our approach, consider a single-phase two-dimensional (2D) numerical reservoir model with a permeability field given in Fig. 1. The model has 16,384 grid blocks, and therefore 16,384 (pressure) states, one injector and four producers (we will explain this example in more detail in Test Case 1 below). Figure 2 displays the first 130 diagonal entries of the balanced Gramians, i.e. the Hankel singular values, for this reservoir system. Note that the y-axis is represented on a logarithmic scale, and that the smallest
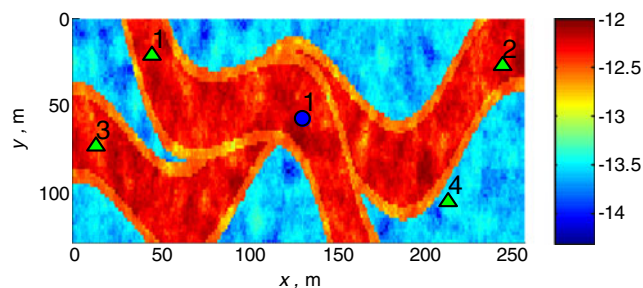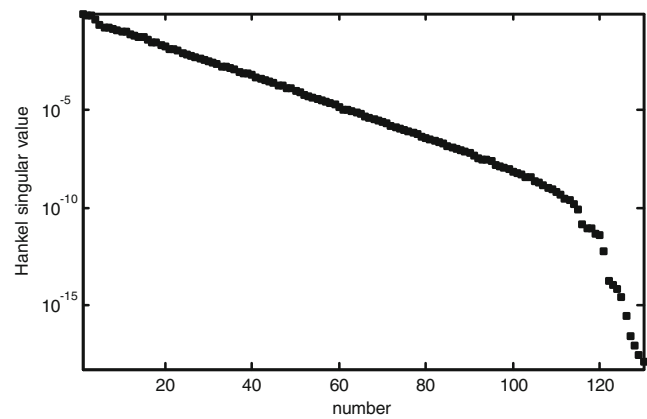

Fig. 2 Hankel singular values for Test Case (1)

singular values of the system are smaller than 10^{-16}, i.e. machine precision in Matlab. The sum of the first 130 squared Hankel singular values equals 1.9977 while the sum of the remaining 16,254 squared singular values, which could not be computed, has an upper bound of only 1.6254×10^{-12}. For practical purposes, the sum of the 130 computed squared singular values can therefore be used as an accurate measure of E_{HSV}. Given that Hankel singular values offer a measure of the degree of the combined controllability and observability of linear combinations of states, the rapid decline in the magnitude of these values implies that the number of combinations required to describe the input/output behavior of the system is much smaller than 16,384. Accordingly, the dominant reservoir dynamics 'lives' in a state-space of much smaller dimensions than suggested by the number of grid blocks.

To visualize the controllability/observability, we consider 'directions' $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_6$, corresponding to the six largest HSVs of the system. Patterns 1 to 6 in Fig. 3 show the directions mapped onto the computational grid. Because each component of the state represents the pressure in a grid block, this mapping allows us to quantify the variation of the controllability and observability of the system over the spatial domain in the form of 'patterns'. For a better visualization, in Fig. 4, we have sorted the grid blocks based on their importance (the magnitude of their non-zero values in the previous figure). The colors therefore represent the 'grid importance'; i.e. the ranked importance of the individual grid block contributions to the system controllability/observability (in terms of the 'input–output energy' E_{HSV}), ordered from high contributions (red) to negligible contributions (blue).

The significance of each mapped pattern is proportional to the magnitude of the related Hankel singular value. The vector sum of all patterns, each weighted
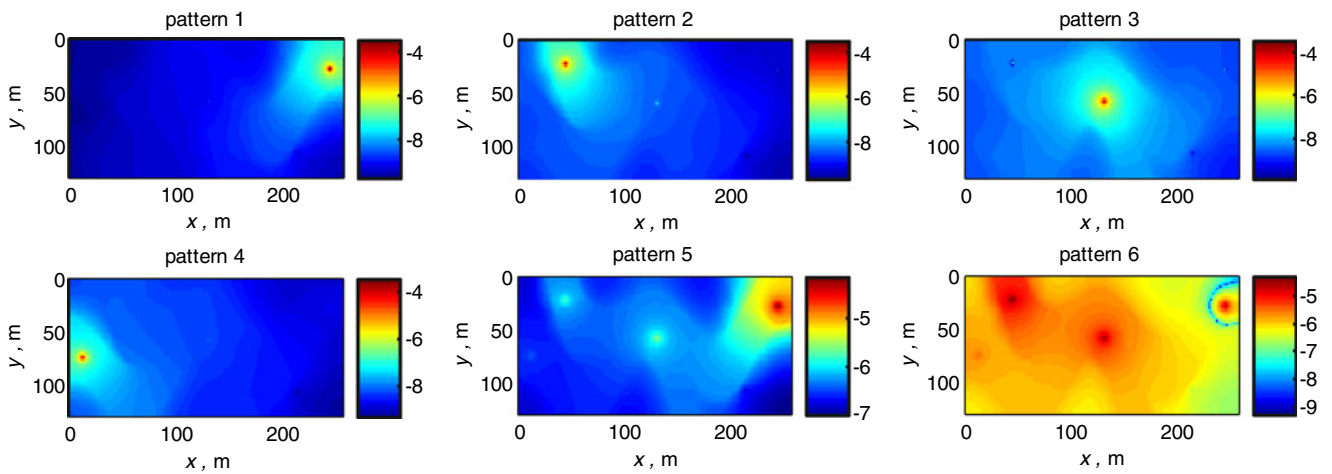

Fig. 1 Log 10 permeability field and injector (*circle*) and produces (*triangles*) locations for Test Case (1)

Fig. 3 Patterns 1 to 6 represent the log10 singular vectors corresponding to the six largest Hankel singular values of the system in Test Case (1)

with its corresponding singular value, therefore gives the 'dominant' pattern that represents the spatial variation of the combined controllability and observability of the system. However, because of the rapid decay of the singular values, only a few patterns related to Σ_1 in Eq. 24 need to be taken into account to accurately capture this combined controllability and observability. Accordingly, the dominant pattern is given by

$$\bar{\mathbf{u}} = \sum_{i=1}^{r} \frac{\sigma_i}{\sigma_1} \mathbf{u}_i. \tag{26}$$

We can select the number of patterns that needs to be retained in the computations, based on the definition of the total energy (Eq. 25). For Test Case (1), the ne-

glected energy for 10, 30 and 60 Hankel singular vectors is 6.8e-3, 5.98e-4, and 1.13e-10 times the (very tight) upper bound to the total energy (E_{HSV}), respectively. Here we choose to use the weighted sum of the first 60 patterns. Figure 5 presents the resulting dominant pattern, both in terms of the log10 singular values and as grid block importance map. It is clear from this figure that, for the given example, the most controllable and observable regions (red) are in the vicinity of the wells and in high-permeable areas connected and close to the wells. This observation is also in line with earlier results from [43]. Note that the area in the vicinity of producer 4 is less controllable/observable, compared to other producers, because it is located in a low-permeable part of the reservoir.
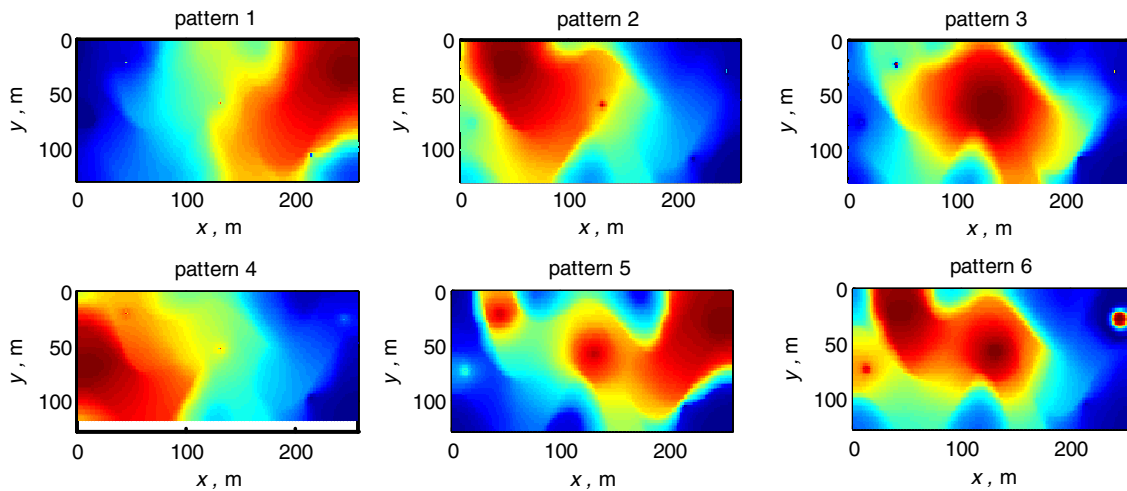

Fig. 4 Grid block importance maps. Visualization of patterns 1 to 6 in Test Case (1) in terms of grid importance, i.e. ranked grid block contribution to E_{HSV}. *Red* High, *blue* low
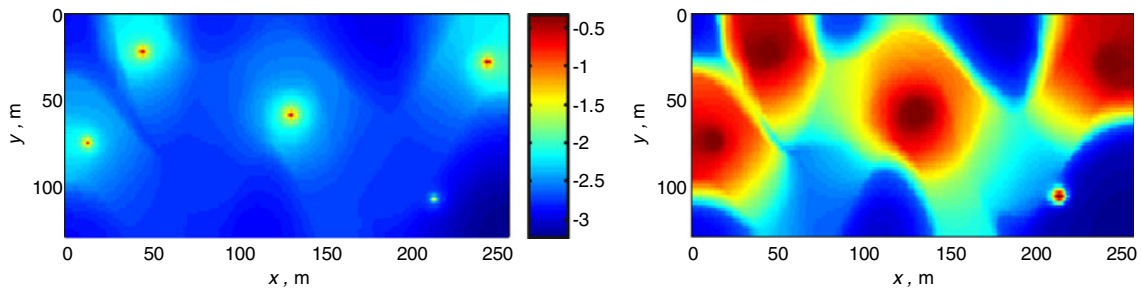
Fig. 5 Visualization of the dominant pattern for controllability and observability variation over the spatial domain for Test Case (1); the Gramians were obtained by LCRF-ADI algorithm. *Left* Weighted sum of the first sixty log10 singular vectors. *Right* Weighted sum of the first sixty grid block importance maps (*red* high importance, *blue* low importance)

This method to rank the grid block contributions to the observability and controllability properties of a reservoir model forms the basis for our CRSC algorithm. The basic idea is to perform grid coarsening only in the weakly controllable and observable areas, i.e. in those areas that have the least effect on the input/output behavior of the system. We note that after the coarsening has been performed we can simply re-compute the upper bound to the total energy E_{HSV} such that it is possible to quantify the effect of the upscaling in terms of system controllability/observability without performing any reservoir simulations.

3.2 CRSC algorithm

Algorithm

The CRSC method can be implemented using direct selective coarsening, or using initial uniform coarsen-

ing followed by selective refinement. In case of direct selective coarsening (see Fig. 6a) we take the following steps:

Algorithm 1
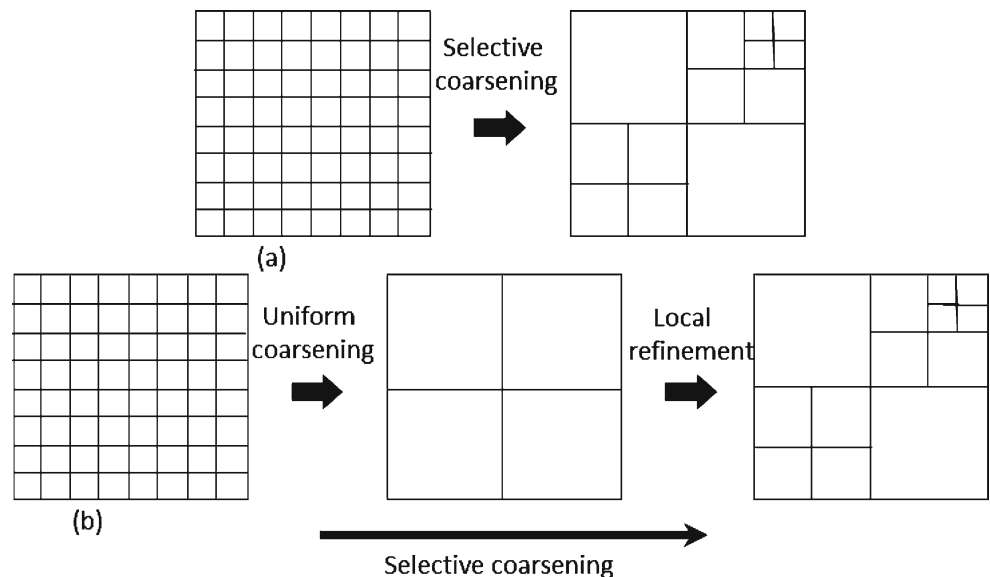
1.1   Spatially quantify the combined controllability and observability of the system based on the fine model.
1.2   Selectively coarsen the domain in the areas of the lowest grid importance, i.e. the areas that contribute least to the combined controllability and observability.

In case of selective refinement after uniform coarsening (see Fig. 6b) we take the following steps:

Algorithm 2a

2a.1   Perform a uniform coarsening with a local up-scaling technique.

Fig. 6 Two strategies to implement CRSC; **a** (*Top*) direct selective coarsening. **b** (*Bottom*) initial uniform coarsening and subsequent selective refinement

2a.2 Spatially quantify the combined controllability and observability of the system based on the *fine* model.
2a.3 Selectively refine the domain in the areas of the highest grid importance, i.e. the areas that contribute most to the combined controllability and observability.

Alternatively, we may modify step 2a.2 to enable a faster but potentially less accurate implementation:

Algorithm 2b

2b.1 Perform a uniform coarsening with a local upscaling technique.
2b.2 Spatially quantify the combined controllability and observability of the system based on the *coarse* model.
2b.3 Selectively refine the domain in the areas of the highest grid importance, i.e. the areas that contribute most to the combined controllability and observability.

In the following, we explain Algorithms 2a and 2b in more detail:

Step (2a/b.1)

We choose an upscaling technique to generate a coarse-scale model. A fast and simple upscaling method is a so-called local upscaling technique, in which coarse-scale parameters are calculated from local flow calculations subject to some generic boundary conditions [40]. An alternative technique is a local–global approach, in which the generic boundary conditions are only used to obtain an initial coarse-scale flow solution. The interpolation of the coarse-scale solution then gives more accurate local boundary conditions to find the new coarse-scale parameters and the process is iterated until the solution converges [7, 13].

Step (2a.2)

To spatially quantify the most controllable/observable patterns of the original fine-scale system, we first need to compute the balanced Gramians of the system by solving Lyapunov equations (Eqs. 20 and 21). The balanced Gramians may then also be used to compute a measure of the system observability/controllability in the form of a tight upper bound to the 'input–output energy' E_{HSV}. An extensive overview of methods to compute exact and approximate solutions to the Lyapunov equations can be found in [2, 25] and the references therein.

We tried three different methods to compute the Gramians. The first method is to use direct algorithms for small dense Lyapunov equations such as the Bartels–Stewart technique [4] and Hammarling's method [16], which are already implemented in Matlab as functions *lyap* and *lyapchol*. A shortcoming is, however, a very high computational time and memory requirement for large systems (with an order of higher than 10^4 cells).

The second option is to use a low-rank iterative approximation of the Gramians. This method has been implemented in LYAPACK [29], which is a Matlab toolbox for the solution of large-scale problems in control theory. It uses iterative algorithms and it is intended for solving large and sparse Lyapunov equations. In many cases like in our application, the eigenvalues of the Gramians are decaying very fast, indicating that there exist accurate low-rank approximations. Let

$$\mathbf{W}_C = \mathbf{L}_C \mathbf{L}_C^T, \tag{27}$$

and

$$\mathbf{W}_O = \mathbf{L}_O \mathbf{L}_O^T, \tag{28}$$

where $\mathbf{L}_C$ and $\mathbf{L}_O$ are the Cholesky factors of the controllability and the observability Gramians, respectively. For large systems we replace the full-rank matrices $\mathbf{L}_C$ and $\mathbf{L}_O$ by a low-rank approximation of Cholesky factors (LRCFs) $\mathbf{L}_C^r$ and $\mathbf{L}_O^r$, where r is the lower rank related to the order of the controllable and observable subspaces. For instance, for Test Case 1 with $n = 16,384$, Fig. 2 suggests that r could be around 60. Consequently, we can store $n \times r$ matrices $\mathbf{L}_C^r$ and $\mathbf{L}_O^r$ instead of the dense $n \times n$ matrices $\mathbf{W}_C$ and $\mathbf{W}_O$. Therefore, in addition to reducing the computational cost, we reduce the memory requirements. The use of LRCFs together with alternating direction-implicit (ADI) iteration algorithm in LYAPACK leads to an efficient technique that can approximate the system Gramians nearly up to the machine precision. The fast and reliable LRCF-ADI approach, therefore, will be used to compute the Gramians in Test Cases 1 and 2 below. For a detailed implementation of this method see [29].

The third method to approximate the system Gramians is based on (balanced) proper orthogonal decomposition ((B)POD). Several authors have investigated the idea of combining the balanced truncation and POD including Moor [26], Willcox and Peraire [42], Lall et al. [23], Antoulas [2], and Rowely [31]. To compute the POD modes of a system for a particular set of well and boundary conditions, one has to perform a fine-scale

simulation and take l snapshots at different points in time to form the snapshot matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_l \end{bmatrix}. \tag{29}$$

A singular value decomposition (SVD) of $\mathbf{X}$ gives $\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, and $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values σ of the system. Moreover, the continuous-time adjoint (dual) model of a linear system is defined as

$$\dot{\mathbf{z}} = \mathbf{A}_c^T \mathbf{z} + \mathbf{C}_c^T \mathbf{u}_a, \tag{30}$$

where $\mathbf{A}_c$ is the continuous-time system matrix, $\mathbf{C}_c$ is the continuous-time output matrix and $\mathbf{u}_a$ represents the adjoint input. It is known that the POD modes of the impulse response of the forward (adjoint) system are equivalent to the dominant eigenvalues of the controllability (observability) Gramian of the system, and accordingly $\mathbf{W}_C \simeq \mathbf{X}\mathbf{X}^T$ and $\mathbf{W}_O \simeq \mathbf{Z}\mathbf{Z}^T$. Therefore, computing the BPOD approximate Gramians involves the simulation of the forward and the adjoint systems, and storing several state and adjoint state snapshots in matrices $\mathbf{X}$ and $\mathbf{Z}$, respectively. Subsequently, from the solution of the SVD problem $\mathbf{Z}^T\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, we can obtain the balanced transformation matrix as

$$\mathrm{T} = \mathbf{X}\mathbf{V}\boldsymbol{\Sigma}^{-1/2}, \tag{31}$$

where the diagonal entries of $\boldsymbol{\Sigma}$ are the Hankel singular values of the system.

Note that if the inputs and outputs are located in the same spatial locations (well positions), we have $\mathbf{B} = \mathbf{C}^T$ and therefore $\mathbf{W}_C = \mathbf{W}_O^T$, in which case we only need to perform POD and no BPOD. Using (B)POD we avoid the expensive step of solving the Lyapunov equations for large systems. However, this comes at the cost of having to perform forward (and adjoint) fine-scale simulations to generate the snapshots required for the (B)POD method. The Hankel singular values and the corresponding directions are then used to quantify the variation of the system's controllability and observability properties over the spatial domain. Figure 7 depicts the dominant pattern obtained from the approximate BPOD Gramians. The resulted mapping is very close to the one in Fig. 5, which was obtained by using the LCRF-ADI technique to calculate the Gramians.

Before finishing step 2a.2, the coarse-scale grid blocks that are located in the areas corresponding to highly controllable/observable states (red areas in Fig. 5 or Fig. 7), are flagged to be refined in the next step. A grid block is flagged if the following condition holds:

$$\frac{\bar{u}_i}{\|\bar{\mathbf{u}}\|} > \alpha, \tag{32}$$

where $0 \leq \alpha \leq 1$ is a threshold value. A zero threshold value means that all the grid blocks should be flagged and the corresponding coarse-scale grid blocks need to be refined, while a unit threshold value means no refinement. From our experience and for the following examples, we found out that setting a threshold value of about 0.005 produced accurate coarse-scale results, while significantly reducing the number of grid blocks. However, the optimal choice of the threshold value is still an open question.

Step (2b.2) An alternative and faster, but potentially less accurate, approach is to use a approximated coarse-scale (but still relatively fine) model to compute the balanced Gramians and perform this step. Steps 2b.2 and 2b.3, in that case, need to be iterated to get more accurate approximation of the controllable/observable subspaces. This approach is particularly attractive for very large fine-scale models, for which a direct calculation of the fine-scale Gramians is infeasible.

Step (2a/b.3) The last step is to refine the flagged coarse-scale grid blocks to finer ones, or even back to the initial fine-scale grid blocks. The result is a system
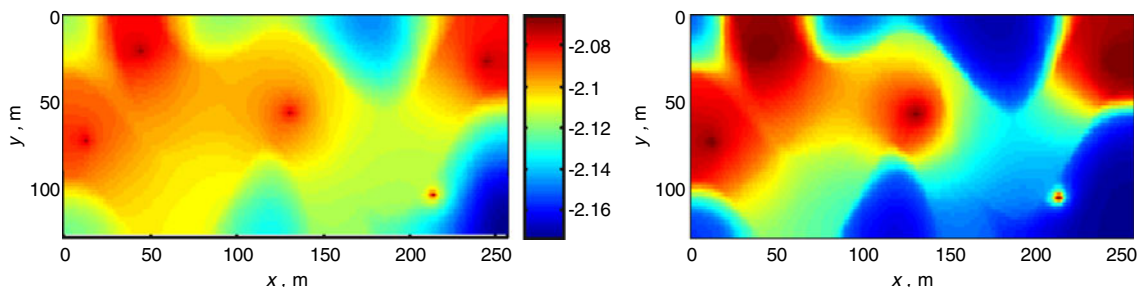

Fig. 7 Visualization of the dominant pattern for controllability and observability variation over the spatial domain for Test Case (1); the Gramians were obtained using BPOD. *Left* Weighted sum of the first eight singular vectors. *Right* Weighted sum of the first eight grid block importance maps (*red* high importance, *blue* low importance)
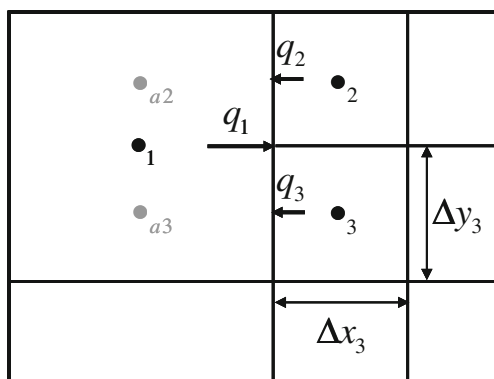
Fig. 8 Non-uniform grid blocks

Fig. 9 Prescribed well bottom-hole pressures for Test Case (1)

of non-uniform grid blocks as depicted in Fig. 8. To discretize partial differential flow equations over non-matching grid blocks, we choose a cell-centered finite volume technique. The interface fluxes for each control volume are calculated by using a two-point flux approximation method [1, 12]. Therefore, the flux between grid blocks 2 and 1 in Fig. 8 is given by

$$q_2 = -\left(\frac{1}{\mu}\frac{2K_2K_1}{\Delta x_2 K_1 + \Delta x_1 K_2}\Delta y_2\right)(p_2 - p_{a2}), \qquad (33)$$

where K_1 and K_2 are the permeability values of grid blocks 1 and 2, p_2 and p_{a2} are pressure values at points 2 and $a2$, μ is the fluid viscosity, and Δx and Δy are the grid block dimensions. One issue here is to calculate the pressure values at the auxiliary points like $a2$ and $a3$. Although there are various ways to do so, for simplicity we assume that they are equal to the average pressure for the entire grid block. However, to maintain the accuracy around the interface, we require that each grid block may only be refined once in each direction, i.e. at each refinement level, a grid block may contain up to four sub grid blocks. More accurate pressure calculations at the auxiliary points can be found in e.g. [12, 13, 15, 22]. Spatial and temporal discretization of the flow equation over the non-uniform grid blocks leads to a system of equations that can be written in state-space form Eq. 14. After completion of step 2a/b.3 we can recompute a tight upper bound to the 'input–output energy' E_{HSV}, and thus quantity the effect of the coarsening on the system controllability/observability.

4 Numerical results

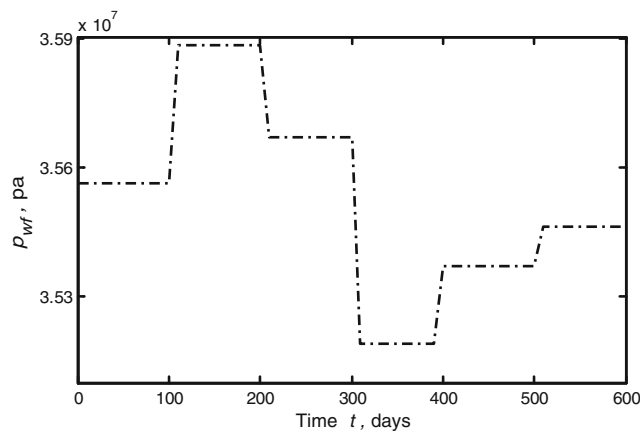The performance of CRSC will be demonstrated with the aid of two numerical test cases. The first one concerns a channelized reservoir with relatively small permeability fluctuations, while the second one concerns a strongly channelized system. In both cases, we assume single-phase flow subject to no-flow boundary conditions. The liquid flow into and out of the reservoir is therefore through the wells. We consider a fixed configuration of the wells which are controlled by a prescribed variable bottom-hole pressure, while production and injection rates are recorded as the output of the system. The bottom-hole pressure is related to the grid block pressure by using a well index which is a function of the grid block geometry and permeability [28]. For Test Case (1), we also compare the results of our non-uniform coarsening method to those of two velocity-based non-uniform upscaling techniques, in which the regions of potentially high fluid velocities are represented with finer grid blocks. To determine those high-velocity regions, we use a flux-based and a streamline-based approach. In the flux-based approach, adapted from Durlofsky et al. [10], first an initial approximate coarse-scale model is constructed. Then the flow rate across each coarse-scale face is approximated as the sum of the flow through the corresponding fine-scale faces. Two coarse-scale grid
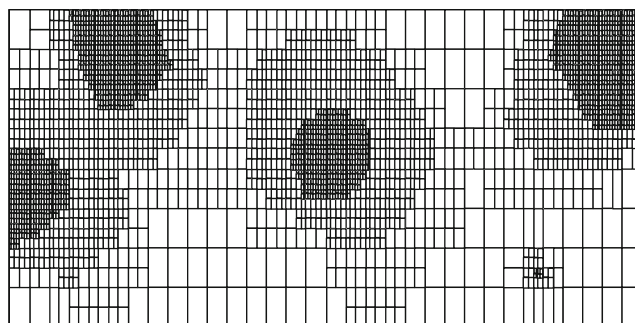

Fig. 10 Selectively coarsened grid for Test Case (1) using the CRSC method (Algorithm 2a)

Table 1 CRSC performance for Test Case (1) using fine-scale Gramians (Algorithm 2a)

Model	Fine-scale	Coarse (32 × 8)	CRSC	Flux-based	Streamline-based	Coarse (128 × 32)
e_q (%)	0	29	0.36	0.36	0.25	13
e_E (%)	0	73	6.4e-3	2.3e-3	1.9e-3	45
n-ratio	1	0.02	0.23	0.23	0.24	0.25

blocks sharing a face are then refined if the approximated flux across that face is larger than a prescribed amount (which is specified as a fraction of the maximum flow rate through any coarse-scale face). Similarly, in the streamline-based approach, regions with high streamline densities are represented with a finer grid (see, e.g. He and Durlofsky [18] and the references therein).

To compare the performance of the non-uniformly coarsened model obtained by the CRSC algorithm with those of the original fine-scale model, the uniformly coarsened model, and the velocity-based non-uniform techniques, we define a relative production error as

$$e_q = \left| \frac{q_{\text{fine}} - q_{\text{coarse}}}{q_{\text{fine}}} \right| \times 100\%, \tag{34}$$

where q is the total cumulative production of the reservoir. We also define an 'energy' error as

$$e_E = \left| \frac{E_{\text{HSV,fine}} - E_{\text{HSV,coarse}}}{E_{\text{HSV,fine}}} \right| \times 100\%. \tag{35}$$

Note that it is possible to compute e_E from the system equations without performing any simulations. Furthermore we define an 'n-ratio' as the ratio of the number of the grid blocks in each model to that of the original fine-scale model. This ratio also relates the number of linear solves and, accordingly, gives a rough estimate of the computational time that is needed to simulate each reservoir model compared to that of the fine-scale one. Clearly, the n-ratio for the fine-scale system is equal to 1.
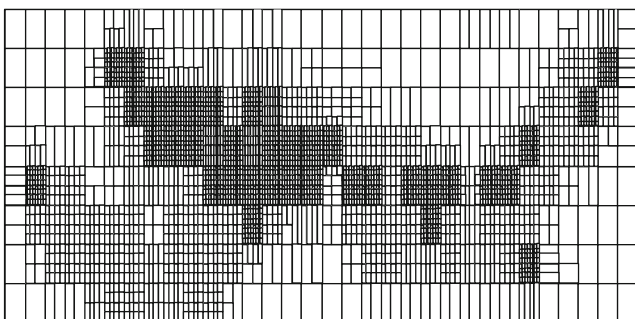
4.1 Test case (1)

The first example is a 2D synthetic reservoir with a permeability field and well locations depicted in Fig. 1. The permeability values vary between 5 and 1,000 mD. There are one injector in the middle and four producers around that, resembling an inverted five-spot pattern. All wells are controlled by prescribed bottom-hole pressures. The time-varying bottom-hole pressure for injector 1 is shown in Fig. 9. Note that the placement of a well in a low-permeable area is not common practice. We did this, nevertheless, for producer 4 to illustrate the controllability/observability variation for different well placements. Moreover, due to the relatively poor resolution of seismic data, in practice the exact location of the channels may be uncertain or unknown.

We chose factors of 8 in each direction to scale up the 256 × 64 fine-scale model to a uniform 32 × 8 coarse model by using a local upscaling technique. The grid refinement for this example was then based on the quantification of the state's combined controllability and observability over the spatial domain using a grid importance map as illustrated in Fig. 5. Here, we used the fine-scale model to compute the balanced Gramians, i.e. we applied Algorithm 2a. The level of refinement for each flagged coarse grid depends on its importance compared to other grid blocks. In red grid blocks, which are located in the most controllable/observable areas, we perform three levels of refinement to return to the original fine-scale model. Figure 10 shows the selectively coarsened grids obtained by CRSC Algorithm 2a. Clearly the refinement around producer 4 is less than around the other wells

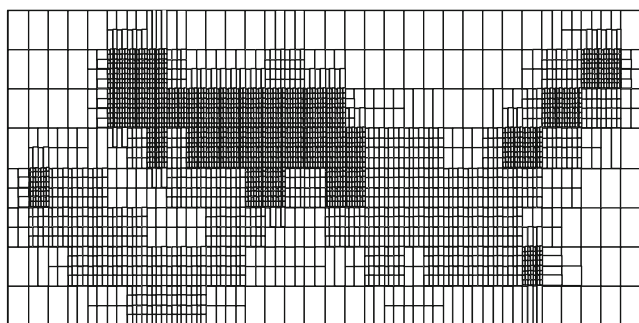**Fig. 11** Selectively coarsened grid for Test Case (1) using the flux-based method
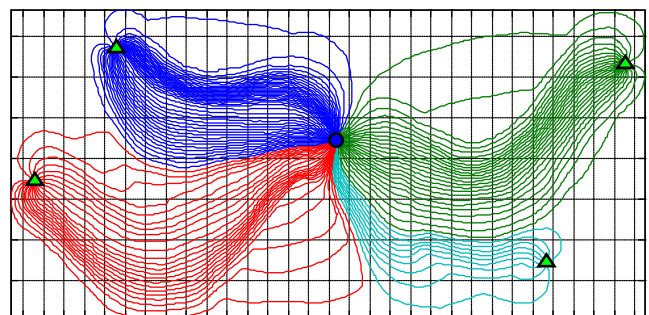
Fig. 12 Streamline groups for Test Case (1)

Fig. 13 Selectively coarsened grid for Test Case (1) using the streamline-based method

Fig. 14 HSV plots for the fine-scale, uniform coarse-scale, and CRSC models in Test Case (1)

because this producer is located in a low-permeable part of the reservoir and it has less effect on the input–output behavior of the system. Table 1 presents the relative error and n-ratio for the simulation of Test Case (1) with different models. To make a fair comparison the boundary conditions for all non-uniform coarse models are the same and the n-ratio is between 0.23 and 0.24. Although the uniform coarse model (32×8) gives a very fast simulation with an n-ratio of 0.02, the CRSC algorithm vastly outperforms the uniformly coarsened grid in terms of accuracy. More interestingly, the CRSC model with 3,794 grid blocks gives a much smaller error than the uniform coarse model (128×32) with 4,096 cells. This illustrates that for this example, in addition to a computational advantage, the selective coarsening is also more accurate than a uniformly upscaled model with the same number of grid blocks.

The performance of the model generated with the CRSC method is comparable to those of the models generated with the flux-based and streamline-based techniques, although the non-uniform grids are slightly different (see Figs. 11, 12 and 13). The streamline-based method gives a slightly better solution, but it also has a slightly larger number of grid blocks (n-ratio of 0.24). Note that the streamline-based technique is more intended for displacement processes, in which it is important to capture high-flow paths to be able to accurately predict the key transport elements such as the breakthrough time of the injected fluid (He and Durlofsky [18]). However, if the inputs rates are changing substantially over the course of the simula-
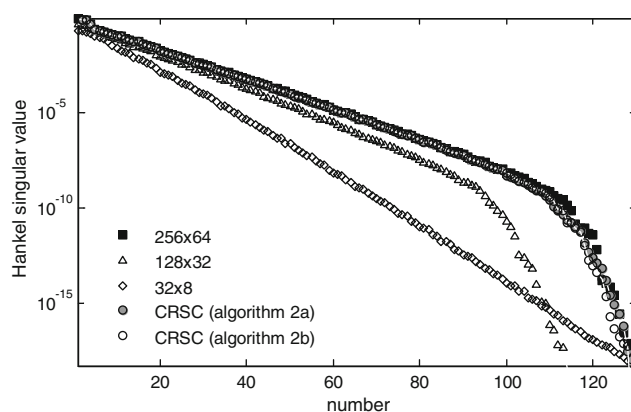
tion, a new streamline pattern may will have to be generated and a new coarsening to be performed. This dependence on the magnitude of the flowrates in the wells is also a feature of the flux-based technique. In contrast, the CRSC method is completely independent of the input rates, and, contrary to the other two non-uniform techniques, it requires no fine- or coarse-scale simulations.

We can also use the Gramians of the initial coarse-scale model for spatial quantification of the controllability/observability, and thus for flagging the most controllable and observable parts of the reservoir (i.e. Algorithm 2b). Table 2 gives the results for Test Case (1) after two iterations. Surprisingly, we observe that, for this example, the errors are comparable to the result obtained from Algorithm 2a, i.e. using the fine-scale Gramians (even without any iteration), while the computational time that is needed to compute the CRSC grid is decreased from 188 s for the fine-scale case to 85 s for the coarse-scale Gramians with one iteration, and to 42 s without any iterations. Note that there is no change in the results after two iterations.

Finally, we compare the HSV plots of all uniform and non-uniform coarse models of Test Case (1) in Fig. 14. As expected, the HSV plots of the CRSC models, using either fine-scale Gramians (Algorithm 2a) or coarse-scale Gramians (Algorithm 2b) are better matched to the HSV plot of the fine-scale model.

Table 2 CRSC performance for Test Case (1) using coarse-scale Gramians (Algorithm 2b)

Iteration	0	1	2
e_q (%)	0.61	0.59	0.59
e_E (%)	1.2e-2	1.1e-2	1.1e-2
Number of grid blocks	3794	3824	3824
Gridding time (s)	42	85	135
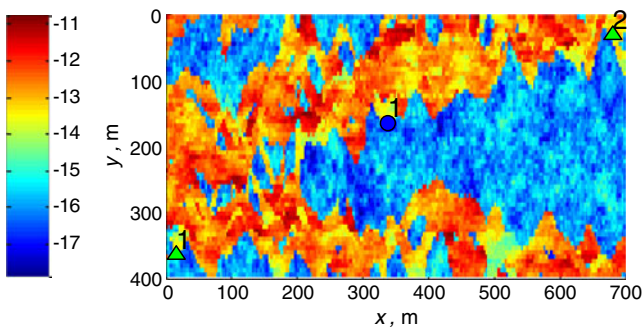
Fig. 15 Log 10 permeability field and injector (*circle*) and produces (*triangles*) locations for Test Case (2)

Fig. 17 Selectively coarsened grid for Test Case (2)

4.2 Test case (2)

The second test case is a 2D reservoir with a channelized permeability field taken from layer 44 of the SPE10 comparative solution project [9], modeled with 220×60 grid blocks. Permeability values for this channelized system are between 0.0001 and 17,000 mD. There are one injection and two production wells with prescribed bottom-hole pressures. The permeability field and the well locations are depicted in Fig. 15. Figure 16 illustrates the dominant pattern obtained from the weighted sum of the Hankel singular vectors corresponding to the Hankel singular values of the system. Recall that the red grid blocks are the most relevant ones in terms of input/output behavior. We observe again that the highly controllable/observable states correspond to the grid blocks in the vicinity of the wells, and of adjacent connected high-permeable zones. Interestingly, the high-permeable areas in the upper left and lower right corners, which are either disconnected or far away from the wells, are weakly controllable/observable.

We scaled up the fine-scale model to a 55×15 coarse model using a local upscaling technique. Subsequent

refinement of the grid blocks corresponding to the most controllable/observable parts of Fig. 16 leads to a pattern of selectively coarsened grid blocks shown in Fig. 17. The original fine-scale grid, the uniformly coarsened grids (55×15) and (110×30), and the CRSC grids were used to simulate the flow behavior in the reservoir. Table 3 gives the simulation results for different models in terms of cumulative production error and the *n*-ratio. As in Test Case 1, it turns out that the CRSC method appropriately reproduces the fine-scale results, and clearly outperforms both uniformly coarsened models obtained with a local upscaling technique. Note that in highly channelized cases local techniques are well known not to resolve the permeability connectivities between different coarse blocks and, therefore, to result in large upscaling errors.

4.3 Complexity analysis

In order to obtain a rough approximation of the computational efficiency of CRSC algorithm, we investigated the computation of the most expensive operations. This includes the computation of the Gramians and the
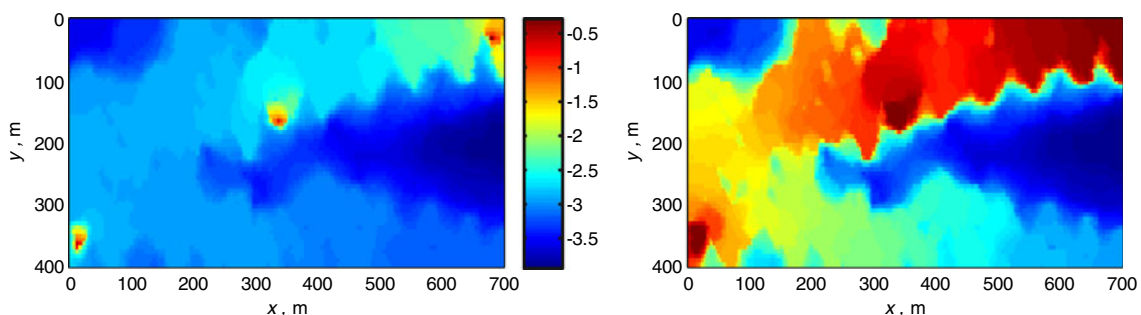

Fig. 16 Visualization of the dominant pattern for controllability and observability variation over the spatial domain for Test Case (2). *Left* Weighted sum of the first 52 log10 singular vectors.
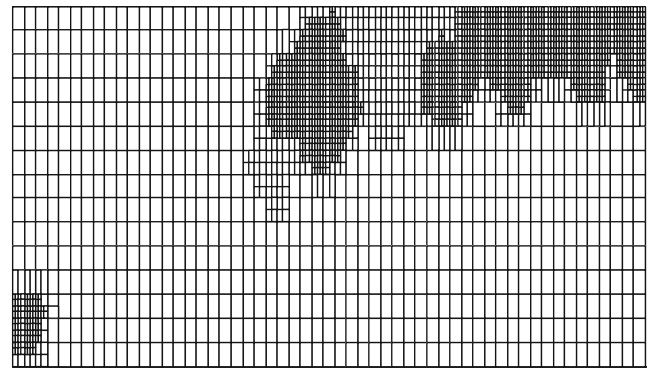
Right Weighted sum of the first 52 grid block importance maps (*red* high importance, *blue* low importance)

Table 3 CRSC performance for Test Case (2)

Model	Fine	Coarse (55×15)	CRSC	Coarse (110×30)
e_q (%)	0	76	0.08	39
e_E (%)	0	82	3.6e-3	43
n-ratio	1	0.06	0.20	0.25

calculation of the Hankel singular values (balancing). We restricted our analysis to single-core implementations, and did not consider ultra large reservoir models (with typically more than 10^6 cells) in which the flow equations need to be solved on multiple cores.

The balancing step requires a singular value decomposition of operation of order n^3. However, in practice, we only perform it for the first ℓ largest singular values and therefore the computational overhead is in the order of ℓ^3, where $\ell \ll n$. This makes the implementation of the SVD on very large matrices still possible. Consequently, the limiting part in the application of CRSC algorithm to large systems is the computation of the Gramians, i.e. the solution of Lyapunov equations. In case of direct Lyapunov solvers, e.g. Stewarts' and Hammarling's Lyapunov solvers, the solver requires operations in the order of n^3 and storage in the order of n^2. The high storage requirement arises because, in addition to producing dense Gramians, these methods are based on Schur decomposition of the sparse system matrix which produces dense matrices. Consequently, these methods are only applicable to systems with an order of $<10^4$ which disqualifies them for realistic reservoir simulation applications.

The second approach is the LRCF-ADI method, for which the operation requirement is in the order of kn where k is the number of columns that are used for the computation. Moreover, the storage cost for this method is much smaller (also in the order of kn). The reason is that here we work directly with low-rank Cholesky factors of the Gramians, compared to direct methods, in which we need to store the full-rank dense Gramians (in the order of n^2). Therefore, this is the method that can be also applied to very large systems. In particular, for large systems in which the input and output matrices **B** and **C** have a low rank, i.e. for small numbers of inputs and outputs, the eigenvalues of the Gramians are decaying very fast, indicating that there exist accurate low-rank approximations. For instance, Fig. 1 shows that the Hankel singular values for Test case (1) are rapidly decreasing. This means that for Test case (1) ($n = 16,384$), k can be chosen equal to 60 or even less (note that the y-axis is represented on a logarithmic scale). Therefore, the use of the LRCF-ADI method for this example can drastically reduce the computational time and storage requirement, compared to direct methods.

The last approach was to approximate the Gramians from the method of snapshots or BPOD. In this case the computational overhead is in the order of l^3, where $l \ll n$ is the number of snapshots. For Test case (1), the number of snapshots is equal to 61. Therefore, this method is also applicable to large systems. However, it requires also the solution of the flow equation over the original fine-scale model and the associated adjoint model with operations in the order of n. For a more detailed complexity analysis of different methods for computation of the Gramians, the reader is referred to [2, 25].

Finally, the grid blocks obtained by the CRSC algorithm are used to solve the flow equation. We assume that the flow equation requires the solution of a system of n equation in each time step, where n is the total number of grid blocks. Note that for large systems it is more efficient to use an iterative method to solve this system of equations. For the pressure solution over the CRSC grid, an extra operation is required to minimize the effective bandwidth. However, it can be neglected particularly for larger systems. The other steps in the simulation are similar for the uniform and CRSC grid systems. This means that for both systems the simulation cost is roughly related to $O(n)$ for an iterative solver.

In summary, for large systems, the CRSC grid can be obtained efficiently by using approximation methods to compute the Gramians, applying a proper grid block numbering, optimizing the bandwidth of the system matrix, and using a proper unstructured data approach to optimize the grid data storing and loading during the simulation.

Note that CRSC grid blocks need to be obtained only once and in an offline part of the simulation, whereas simulation of the flow equation over CRSC grid blocks might be performed several times in iterative procedures. This holds in particular for flooding optimization for a given well configuration. In that case multiple control settings have to be simulated for a single or multiple reservoir models with given model parameters. The method may also be attractive for iterative computer-assisted history matching, although in that case strong changes to the model parameters may require a repetition of the upscaling procedure. In this light we note that for single-phase flow the identifiability of parameter values is closely related to the observability of

state variables [43], which implies that strong updates of parameter values outside the observable region, and therefore the need to repeat the upscaling, are unlikely. However, the identifiability of parameter values in multi-phase flow, after breakthrough in the wells, is more complex, and the computational advantage of CRSC for history matching of multi-phase flow models therefore needs further investigation.

4.4 Discussion

This research is primarily concerned with the system-theoretical aspects of upscaling in reservoir simulation. The main objective is to exploit information in high-resolution models that is relevant for control and measurement for a given configuration of wells. In [42], we demonstrated that the observability of reservoir pressures from output in the wells is very limited and that therefore also the identifiability of reservoir parameters from production data is very limited. At the same time there is a very limited controllability of the reservoir pressures by manipulating the inputs in the wells. Pressures and their spatial derivatives, i.e. velocities, determine the propagation of fluids through the reservoir. For a system with inputs and outputs in the same wells, the (few) identifiable parameter patterns correspond just to the (few) controllable state patterns that are available to influence the fluid flow in the reservoir [20, 21, 43]. The CRU and CRSC methods therefore aim at maintaining precisely those aspects of the reservoir system model that are essential for the 'closed-loop' control of reservoir fluid flow based on production measurements. The computational efficiency of the presented algorithms has not been optimized. In particular, for the sake of simplicity, we used simple gridding and discretization techniques, and we implemented them in MATLAB. Alternative gridding and discretization strategies might improve the CRSC performance in terms of both accuracy and computational efficiency. The CRSC method can also conceptually be extended to 3D applications.

The CRSC method as presented here is a single-phase upscaling technique. The performance of the resulting CRSC grids for two-phase flow applications needs further research. Furthermore, the CRSC algorithm can, in theory, also be extended to two-phase flow simulations. However, for nonlinear two-phase flow cases, we need to either linearize the system, or compute the controllability and observability Gramians empirically (for an empirical calculation of the Gramians, see e.g. [31, 42]). Another important issue in two-phase flow is to deal with moving saturation fronts. Since the saturations are only controllable along the front

and only observable after water breakthrough in the wells [39], the performance of CRSC algorithm would be improved by adding grid adaptation to resolve the strongly controllable/observable areas along the moving saturation fronts. Further research is required to evaluate the computational benefits of CRSC approach in multi-phase flow applications. Another possible solution to treat saturation changes in a two-phase flow simulation might be applying a multi-scale framework, in which the CRSC grid blocks are only used to solve the pressure equation while the saturation equation is solved over the fine-scale grid blocks.

5 Conclusions

Based on the theory and the numerical examples discussed above, we reached the following conclusions:

1. Computation of the Hankel singular vectors of a reservoir model and mapping them on the reservoir grid allows for spatial quantification of the combined controllability and observability properties of the model.
2. An improved visualization is obtained by ranking the grid blocks according to their relative contribution to the 'input–output energy' E_{HSV}.
3. For a given configuration of wells, the most controllable/observable areas appear to be in the vicinity of the wells and in high-permeable areas close to and connected to the wells.
4. CRSC can either be achieved by direct selective coarsening or by initial uniform coarsening and subsequent selective refinement in areas of highest combined controllability and observability.
5. For the case of initial uniform coarsening followed by selective refinement it is possible to compute the required system properties using the fine grid or the coarse grid. The latter is faster but potentially more inaccurate. In our test cases no significant deterioration of accuracy was observed.
6. Performing the computation of the Hankel singular vectors using approximation and iteration methods, such as LRCF-ADI algorithm and the 'method of snapshots', allows for application of the CRSC algorithm to very large models, say models up to 10^6 gridblocks which can still be run on a single core. Application to ultra large models, which require multi-core implementations, is currently out of scope.
7. In the test cases discussed in this paper, the CRSC algorithm resulted in models that accurately

reproduce the flow response of the fine-scale models for time-varying inputs.

8. In addition to giving a computational advantage, the selective coarsening gives also more accurate results than a uniformly upscaled model with the same number of grid blocks.

9. In contrast to conventional flow-based coarsening techniques our method is independent from the specific flow rates or pressures imposed at the wells.

10. In addition it is possible to compute a tight upperbound to the 'input–output energy' E_{HSV}, for the original and the upscaled models which is independent of specific well flow rates.

References

1. Aavatsmark, I.: An introduction to multipoint flux approximations for quadrilateral grids. Comput. Geosci. **6**, 405–432 (2002)
2. Antoulas, A.C.: Approximation of Large-scale Dynamical Systems. SIAM, Philadelphia (2005)
3. Aziz, K., Settari, A.: Petroleum Reservoir Simulation. Applied Science, London (1979)
4. Bartels, R.H., Stewart, G.W.: Solution of the matrix equation AX + XB = C. Commun. ACM **15**(9), 820–826 (1972)
5. Berger, M.J., Colella, P.: Local adaptive mesh refinement for shock hydrodynamics. J. Comput. Phys. **82**, 64–84 (1989)
6. Cardoso, M.A., Durlofsky, L.J., Sarma, P.: Development and application of reduced-order modeling procedures for subsurface flow simulation. Int. J. Numer. Methods Eng. **77**(9), 1322–1350 (2008)
7. Chen, Y., Durlofsky, L.J., Gerritsen, M.G., Wen, X.H.: A coupled local–global upscaling approach for simulating flow in highly heterogeneous formations. Adv. Water Resour. **26**, 1041–1060 (2003)
8. Chen, Y., Oliver, D.S.: Ensemble-based closed-loop production optimization on the Brugge field. SPE J. **14**(4), 634–645 (2009)
9. Christie, M.A., Blunt, M.J.: Tenth SPE comparative solution project: a comparison of upscaling techniques. SPE Reserv. Evalu. Eng. **4**(4), 308–317 (2001)
10. Durlofsky, L.J., Jones, R.C., Milliken, W.J.: A nonuniform coarsening approach for the scale up of displacement processes in heterogeneous porous media. Adv. Water Resour. **20**, 335–347 (1997)
11. Durlofsky, L.J.: Upscaling and gridding of fine scale geological models for flow simulation. In: Paper Presented at the 8th International Forum on Reservoir Simulation, Stresa, Italy, 20–24 Jun 2005
12. Edwards, M.G.: Elimination of adaptive grid interface errors in the discrete cell centered pressure equation. J. Comp. Phys. **126**, 356–372 (1996)
13. Gerritsen, M.G., Lambers, J.V.: Integration of local–global upscaling and grid adaptivity for simulation of subsurface flow in heterogeneous formations. Comput. Geosci. **12**, 193–208 (2008)
14. Gildin, E., Klie, H., Rodriguez, A., Wheeler, M.F., Bishop, R.H.: Development of low-order controllers for high-order reservoir models and smart wells. In: Paper SPE 102214 Presented at the SPE Annual Technical Conference and Exhibition, San Antonio, USA, 24–27 Sept 2006
15. Ham, F.E., Lien, F.S., Strong, A.B.: A Cartesian grid method with transient anisotropic adaptation. J. Comp. Phys. **179**, 469–494 (2002)
16. Hammarling, S.: Numerical solution of the stable, nonnegative definite Lyapunov equation. IMA J. Numer. Anal. **2**, 303–323 (1982)
17. Hanzon, B.: The area enclosed by the (oriented) Nyquist diagram and the Hilbert–Schmidt–Hankel Norm of a linear system. IEEE Trans. Automat. Contr. **37**(6), 835–839 (1992)
18. He, C., Durlofsky, L.J.: Structured flow-based gridding and upscaling for modeling subsurface flow. Adv. Water Resour. **29**, 1876–1892 (2006)
19. Heijn, T., Markovinović, R., Jansen, J.D.: Generation of low-order reservoir models using system-theoretical concepts. SPE J. **9**(2), 202–218 (2004)
20. Jansen, J.D., Bosgra, O.H., van den Hof, P.M.J.: Model-based control of multiphase flow in subsurface oil reservoirs. J. Process Control **18**(9), 846–855 (2008)
21. Jansen, J.D., Douma, S.G., Brouwer, D.R., Van den Hof, P.M.J., Bosgra, O.H., Heemink, A.W.: Closed-loop reservoir management. In: Paper SPE 119098 Presented at the SPE Reservoir Simulation Symposium, The Woodlands, USA, 2–4 Feb 2009
22. Khattri, S.K., Fladmark, G.E., Dahle, H.K.: Control volume finite difference on adaptive meshes. In: Paper Presented at the 16th International Conference on Domain Decomposition Methods, New York, USA, 12–15 Jan 2005
23. Lall, S., Marsden, J.E., Glavaski, S.: Empirical model reduction of controlled nonlinear systems. In: Proceedings of the IFAC World Congress, vol. F, pp. 473–478 (1999)
24. Markovinoviæ, R., Geurtsen, E.L., Heijn, T., Jansen, J.D.: Generation of low-order reservoir models using POD, empirical Gramians and subspace identification. Proc. 8th European Conference on the Mathematics of Oil Recovery (ECMOR VIII), Freiberg, Germany, 3–6 Sept 2002
25. Markovinović, R.: System-theoretical model reduction for reservoir simulation and optimization. PhD dissertation, Delft University of Technology (2009)
26. Moore, B.C.: Principal component analysis in linear systems: controllability, observability, and model reduction. IEEE Trans. Automat. Contr. **26**(1) (1981)
27. Nævdal, G., Brouwer, D.R., Jansen, J.-D.: Waterflooding using closed-loop control. Comput. Geosci. **10**(1), 37–60 (2006)

28. Peaceman, D.W.: Interpretation of well-block pressures in numerical reservoir simulation with nonsquare grid blocks and anisotropic permeability. SPE J. **23**(3), 531–543 (1983)
29. Penzl, T.: LYAPACK: A MATLAB Toolbox for Large Lyapunov and Riccati Equation, Model Reduction Problems, and Linear-quadratic Optimal Control Problems. SFB 393 Fakultät für Mathematik, TU Chemnitz (1999)
30. Renard, Ph., Marsily, G.: Calculating effective permeability: a review. Adv. Water Resour. **20**(5–6), 253–278 (1997)
31. Rowely, C.W.: Model reduction for fluids using balanced proper orthogonal decomposition. Int. J. Bifurc. Chaos **15**(3), 997–1013 (2005)
32. Sarma, P., Durlofsky, L.J., Aziz, K., Chen, W.H.: Efficient real-time reservoir management using adjoint-based optimal control and model updating. Comput. Geosci. **10**(1), 3–36 (2006)
33. Skogestad, S., Postlethwaite, I.: Multivariable Feedback Control, Analysis and Design, 2nd edn. Wiley, Chichester (2005)
34. Vakili-Ghahani, S.A.: Control-relevant upscaling. PhD dissertation, Delft University of Technology (2010)
35. Vakili-Ghahani S.A., Markovinović, R., Jansen, J.D.: Upscaling of large reservoir systems by using a control-relevant approach. In: Proc. 11th European Conference on Mathematics in Oil Recovery (ECMOR X1), Bergen, Norway, 8–11 Sept 2008
36. Vakili-Ghahani, S.A., Jansen, J.D.: Control-relevant selective grid coarsening. In: Paper SPE 124581 Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, USA, 4–7 Oct 2009
37. Vakili-Ghahani, S.A., Jansen, J.D.: Control-relevant upscaling. SPE J. **15**(2), 471–479 (2010)
38. Van Doren, J.F.M., Van den Hof, P.M.J., Jansen, J.D., Bosgra, O.H.: Determining identifiable parameterizations for large-scale physical models in reservoir engineering. In: Proc. International Federation for Automatic Control (IFAC) World Congress, Seoul, Korea, 6–11 Jul 2008
39. Van Doren, J.F.M.: Model structure analysis for model-based operation of petroleum reservoirs. PhD dissertation, Delft University of Technology (2010)
40. Warren, J.E., Price, H.S.: Flow in heterogeneous porous media. SPE J. **1**(3), 153–169 (1961)
41. Wen, X.-H., Gomez-Hernandez, J.: Upscaling hydraulic conductivities in heterogeneous media: an overview. J. Hydrol. **183**, ix–xxxii (1996)
42. Willcox, K., Peraire, J.: Balanced model reduction via the proper orthogonal decomposition. AIAA J. **40**(11), 2323–2330 (2002)
43. Zandvliet, M.J., van Doren, J.F.M., Bosgra, O.H., Jansen, J.D., van den Hof, P.M.J.: Controllability, observability and identifiability in single-phase porous media flow. Comput. Geosci. **12**(4), 605–622 (2008)