OPTIMAL CONTROL OF PARTIALLY MISCIBLE TWO-PHASE FLOW WITH APPLICATIONS TO SUBSURFACE CO_2 SEQUESTRATION

MORITZ SIMON* AND MICHAEL ULBRICH†

Abstract. Motivated by applications in subsurface CO_2 sequestration, we investigate constrained optimal control problems with partially miscible two-phase flow in porous media. The objective is, e.g., to maximize the amount of trapped CO_2 in an underground reservoir after a fixed period of CO_2 injection, where the time-dependent injection rates in multiple wells are used as control parameters. We describe the governing two-phase two-component Darcy flow PDE system and formulate the optimal control problem. For the discretization we use a variant of the BOX method, a locally conservative control-volume FE method. The timestep-wise Lagrangian of the control problem is implemented as a variational form in the PDE toolbox Sundance, which is part of the HPC software Trilinos. The resulting MPI parallelized Sundance state and adjoint solvers are linked to the interior point optimization package IPOPT. Finally, we present some numerical results in a heterogeneous model reservoir.

Key words. partially miscible two-phase flow, optimal control, adjoint approach, complementarity condition, control-volume FE method, variational formulation, CO_2 sequestration.

AMS subject classifications. primary 35Q35, 90C90; secondary 65M08.

1. Introduction.

In this paper, we present an adjoint based [19] approach for optimal control problems that are governed by multiphase multicomponent flow in porous media. The concrete application motivating our work is optimal CO_2 sequestration in underground reservoirs: The goal is to maximize the amount of trapped CO_2 after a fixed time interval of injection into a 2D reservoir Ω. To this end, we use a partially miscible two-phase two-component flow model [4]. The discretization (currently in 2D) is done with the aid of the PDE toolbox Sundance [22], which is part of the Trilinos framework [16]. We extended Sundance by the finite volume (FV) and upwinding capabilities required to implement the BOX method [2] and developed an interface between Sundance and the interior point optimization package IPOPT [32]. Extensions to 3D and more complex flows are possible.

For the discretization of the state system PDEs we apply a variant of the BOX method [2], a locally conservative control-volume FE method. For further information on this method see [17, 18]. Other possible numerical methods include streamline diffusion techniques [35] and conservative mixed FE methods [26]. For a survey of suitable FE/FV methods for multiphase multicomponent flows in porous media we refer to [11] and references therein.

Concerning related work on PDE-constrained optimization in reservoir modeling, the main focus has so far not been on CO_2 trapping mechanisms, but rather on inverse history matching, i.e., parameter estimation in petroleum reservoirs (see [25] for an overview), or on the economic aspect to maximize oil production. To this end, optimal control of mixed CO_2–water injection was first considered in [24], using IMPES simulation of a modified black oil model and gradient methods for optimization. For recent approaches to optimal control of enhanced oil recovery (EOR) via water-flooding we refer to the surveys [8, 20] and to the original articles [1, 9, 12, 27, 29, 30, 31, 33].

The optimal control problem we consider here is the optimal sequestration of CO_2 in underground reservois. To obtain a suitable two-phase model, we restrict our attention to the phases CO_2 and water/brine, neglecting oil and further substances. This leads to a partially miscible two-phase flow model. Note that the mentioned control approaches to optimal water-flooding usually use immiscible two-phase or black oil type three-phase flow models. Thus,

*Chair of Mathematical Optimization, Department of Mathematics, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München, Germany (simon@ma.tum.de).

†Chair of Mathematical Optimization, Department of Mathematics, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München, Germany (mulbrich@ma.tum.de).

in our context, the miscibility of CO_2 in brine introduces a further degree of complexity in the state system. We model the switching between saturation and undersaturation of the wetting phase with CO_2 by a (smoothed) complementarity condition.

Concerning our software choices, we intended to stay as flexible as possible. We work with Sundance, which among other advantages supports a direct variational problem formulation [23] that, in particular, supports the direct derivation of the exact discrete adjoint. To achieve this, the timestep-wise Lagrangian of the control problem is formulated as a variational form. We have implemented an interface to the interior point software IPOPT [32], a state-of-the-art nonlinear optimization code that can handle general constraints and provides L-BFGS [36] approximations to the Hessian of the Lagrangian. In [14] several test problems are considered to show that limited-memory BFGS techniques are superior to other comparable algorithms when it comes to EOR history matching with multiphase flow.

The CO_2 injection takes place at wells in a subregion $\Omega_c \subset \Omega$ of the underground reservoir Ω, being purely saturated with saline water (called "brine") at the beginning. The thermodynamical reservoir conditions are chosen such that the injected CO_2 is in supercritical state, so we consider a continuous underground flow model with two phases $\alpha \in \{w, n\}$ and two components $i \in \{1, 2\}$. This two-phase flow is assumed partially miscible in the following sense: On the one hand, the CO_2 component $i = 2$ is expected to dissolve in the water-rich wetting phase $\alpha = w$. On the other hand, the "evaporation" of the water component $i = 1$ into the CO_2-rich nonwetting phase $\alpha = n$ is neglected.

The reservoir medium is assumed rigid with time-invariant functions $\phi = \phi(x)$ and $\mathbf{K} = \mathbf{K}(x)$ for porosity and absolute permeability. Moreover we neglect temporal and spatial variations of the reservoir temperature and assume chemical equilibrium. The most important physical agreement that determines the flow model's nature is the application of Darcy's law in the PDE system; this means we must exclude reservoir regions with high porositiy $\phi \approx 1$, where the fluids rather obey the Navier–Stokes equations and could exhibit turbulences.

The paper is organized as follows: Section 2 introduces the state equations for Darcy based partially miscible two-phase flow, together with primary state variables and control objectives. Section 3 presents the adjoint based optimal control framework, while the applied numerical methods for the state equation are sketched in Section 4. Implementation strategies for the adjoint based framework within the PDE toolbox Sundance will be discussed in Section 5. Finally, Section 6 contains a discussion of optimization results.

2. State equations, variables and control objectives. We consider partially miscible two-phase flow based on Darcy's law. The mass conservation equations for the components $i \in \{1, 2\}$ ($1 =$ water, $2 = CO_2$) within the phases $\alpha \in \{w, n\}$ (wetting and nonwetting) [4] are then given by

$$\phi \partial_t \sum_\alpha \rho_\alpha X_\alpha^i S_\alpha - \sum_\alpha \operatorname{div}\Big[\lambda_\alpha \rho_\alpha X_\alpha^i \mathbf{K}\big(\nabla p_\alpha + \rho_\alpha \mathfrak{g} e_2\big)\Big] - \operatorname{div}\big(D^i \rho_w \nabla X_w^i\big) = q^i. \quad (2.1)$$

Here ρ_α, S_α and p_α denote the phase densities, saturations and pressures, while X_α^i stands for the mass fraction of component i in phase α. Especially, we have $X_n^1 = 0$ and $X_n^2 = 1$ in the partially miscible setting, while we set $X := X_w^2$ so that $X_w^1 = 1 - X$. The first divergence term in (2.1) expresses advection and buoyancy via Darcy's law with phase mobilities λ_α and absolute permeability matrix $\mathbf{K}$, whereas the second divergence term models the slow diffusion in the wetting phase. Next, we explain some modeling agreements.

For phase mobilities λ_α and capillary pressure p_c we use Brooks–Corey models [7]; the phase viscosities are assumed constant. The brine phase is incompressible, but a linear density increase with the CO_2 mass fraction X is taken into account. Slight compressibility of the CO_2 phase is expressed in a linearized density $\rho_n(p_n)$. The nonlinear diffusion coefficient

$D^2 = D^2(\phi, S_w)$ is taken from Section 2.5 of [4], while water–brine diffusion is neglected via $D^1 = 0$. For further details on these models we refer to [28].

As we do not consider thermodynamical heat exchange, the governing PDE system is given by (2.1) for the components water and CO_2. However, the CO_2 component is allowed to dissolve in the wetting phase $\alpha = w$ to some extent. Hence we need three primary state variables $y := (p_n, S_w, X)$ in the flow model:

1. CO_2 pressure: $p_n = p_n(x,t) \quad \longrightarrow \quad p_w = p_n - p_c \quad$ with $\quad p_c = p_c(S_w)$,
2. brine saturation: $S_w = S_w(x,t) \quad \longrightarrow \quad S_n = 1 - S_w \quad$ (natural constraint),
3. CO_2 mass fraction: $X = X(x,t) \quad \longrightarrow \quad X = \varphi(p_n) \quad$ when $\quad S_w(x,t) < 1$.

As an initial condition we state $S_w(x,0) \equiv 1$ (no CO_2 injected); in such single-phase regimes the pressure $p_n = p_w$ exists as a mutual parameter. Any value $X < \varphi(p_n)$ corresponds to $S_w = 1$, while the nonwetting phase $\alpha = n$ can only appear (that means $S_w < 1$) if X reaches the threshold $\varphi(p_n)$. In the presence of the nonwetting phase we always have the relation $X = \varphi(p_n)$. Thus an adaptive switching between the state variables X and S_w is required; we model this by a complementarity condition:

$$a := 1 - S_w \geq 0, \quad b := \varphi(p_n) - X \geq 0 \quad \text{and} \quad ab = 0.$$

This can equivalently be written as $\tilde{g}(a,b) = 0$, where $\tilde{g}$ is a complementarity function, e.g., the Fischer–Burmeister [13] function $\tilde{g}(a,b) = a + b - \sqrt{a^2 + b^2}$. To maintain differentiability, we smooth this nonsmooth equation in our implementation to obtain an algebraic equation $g(p_n, S_w, X) = 0$ that holds pointwise in space and time (see [28] for details).

With the above state variables $y = (p_n, S_w, X)$, connected through the complementarity condition $g(y) = 0$, the flow system (2.1) can be recast in a more compact form: Introducing the scalar and vector valued abbreviations

$$b^1 := \rho_w(1-X)S_w, \quad b^2 := \rho_n(1-S_w) + \rho_w X S_w,$$

$$B^1 := \lambda_w \rho_w (1-X)\mathbf{K}\big[\nabla(p_n - p_c) + \rho_w \mathfrak{g} e_2\big] - D^1 \rho_w \nabla X,$$

$$B^2 := \lambda_w \rho_w X \mathbf{K}\big[\nabla(p_n - p_c) + \rho_w \mathfrak{g} e_2\big] + \lambda_n \rho_n \mathbf{K}\big(\nabla p_n + \rho_n \mathfrak{g} e_2\big) + D^2 \rho_w \nabla X,$$

we obtain an algebro-differential system of the form

$$\phi \partial_t b^i(y) - \operatorname{div} B^i(y, \nabla y, x) = q^i(u) \quad (i = 1,2) \quad \text{and} \quad g(y) = 0 \tag{2.2}$$

in the open space-time domain $\Omega \times]0,T[$. The spatial boundary $\partial\Omega = \Gamma_n \cup \Gamma_d$ is decomposed into Neumann and Dirichlet parts: A no-flux condition $n^T B^i(y, \nabla y, x) = 0$ for $i = 1,2$ is considered on Γ_n (impermeable cap rock). The conditions $y = y_d$ on Γ_d impose standard values without CO_2 injection, also taken as initial values inside the reservoir.

The control parameters $u = u(t)$ only appear in the source terms $q^i = q^i(u)$:

$$q^1 = 0 \quad \text{and} \quad q^2 = \sum_{n=1}^{N} q_n(t)\omega(x - x_n).$$

This means no water is injected, while q^2 models the injection of supercritical CO_2 at $N \in \mathbb{N}$ injection wells with locations x_n in a certain control area $\Omega_c \subset \Omega$; ω models the injection distribution, given, e.g., by a narrow Gaussian normal distribution. In the following, the N injection rates $q_n(t)$ are used as control parameters:

$$u(t) = \big(q_1(t), \ldots, q_N(t)\big), \quad t \in [0,T]. \tag{2.3}$$

We impose box constraints on the above rates and make sure that the total amount of injected CO_2 does not exceed a certain threshold. This leads to control constraints

$$0 \leq q_n(t) \leq a, \quad t \in [0,T], \quad 1 \leq n \leq N \quad \text{and} \quad \frac{1}{T}\sum_{n=1}^{N}\int_0^T q_n(t)\,dt \leq b \tag{2.4}$$

for fixed numbers $a, b > 0$ with $b < Na$ (otherwise the latter constraint is redundant).

Our objective is to maximize the amount of trapped CO_2 after a finite time $T > 0$ of injection. Here we consider the following types of "trapping": CO_2 can be

1. immobilized below a certain residual saturation S_n^r (residual trapping),
2. dissolved in the brine and thereby trapped for the moment (solubility trapping).

The residual saturation $S_n^r = S_n^r(\phi)$ depends on the porosity. Below this threshold, i.e., for saturations $S_n \leq S_n^r$, the nonwetting phase mobility λ_n drops to zero. For a discussion of further trapping objectives we refer to [28].

A general optimization objective concerns the final amount of trapped CO_2 in either of the above variants, weighted by coefficients $\beta_{1,2} \in [0,1]$, which results in an objective functional of the form

$$J(y,u) = \int_\Omega \left(\beta_1 \phi \rho_n S_n \mathbf{1}_{\{S_n \leq S_n^r\}} + \beta_2 \phi \rho_w X S_w\right)\big|_{t=T}\,dx - \gamma R(u), \tag{2.5}$$

where an additional control-dependent penalty term $R(u)$ allows for regularization of the optimization algorithm in terms of a user-defined parameter $\gamma \geq 0$. It is chosen in a weighted H^1–type manner to penalize strong fluctuations in the wells' rates with an additional weighted L^2–type regularization. Explicitly, our regularization is given by

$$R(u) := \frac{q_*^2}{T}\sum_{n=1}^{N}\int_0^T \left[T^2 q_n'(t)^2 + \theta(t) q_n(t)^2\right] dt, \tag{2.6}$$

including a scaling factor q_* ($q_* = 200$ in SI units) and a weight function $\theta : [0,T] \to [0,\infty[$ that introduces certain time preferences. This strictly convex regularization smooths out the optimization algorithm for reasonable choices of the regularization parameter γ and thereby enhances the flexibility and robustness of the optimization interface.

We finally remark that many of the concrete choices made in this section could be replaced by a wide range of other choices without leaving our overall framework.

3. Adjoint based optimal control framework. Provided that the weak state equation $E(y,u) = 0$ has a unique solution $y = y(u)$ for every control u and, moreover, sufficient smoothness of the functionals $E(y,u)$ and $J(y,u)$, the adjoint approach [19] can be used for efficient computation of the gradient of the reduced functional $j(u) := J\big(y(u),u\big)$:

$$j'(u) = E_u(y,u)^*\mu + J_u(y,u),$$

where the adjoint state μ solves the adjoint equation

$$E_y(y,u)^*\mu + J_y(y,u) = 0.$$

In order to calculate $j(u)$ and $j'(u)$, discretizations of both state and adjoint equation must be set up (see [28] for the derivation of the continuous adjoint equation). The numerical methods we apply are implemented such that the discretized adjoint system coincides with the exact discrete adjoint of the discretized state equation ("discretize–then–optimize").

The time horizon $t \in [0,T]$ is discretized into n_t equidistant time steps t_k with step size $\Delta t := t_k - t_{k-1}$. The corresponding vector

$$u \in \mathbb{R}^{Nn_t} \quad \text{with} \quad u_{(n-1)n_t+k} := q_* q_n(t_k), \quad 1 \le k \le n_t, \quad 1 \le n \le N, \tag{3.1}$$

discretizes the control variables (2.3) with the mentioned scaling factor q_*. The control constraints (2.4) in discrete form read

$$0 \le u_j \le q_* a, \quad 1 \le j \le Nn_t \quad \text{and} \quad c(u) := \frac{1}{n_t} \sum_{j=1}^{Nn_t} u_j \le q_* b,$$

while the penalty term (2.6) can be discretized as

$$R(u) \approx \sum_{n=1}^{N} \left\{ n_t \sum_{k=1}^{n_t-1} \left[u_{(n-1)n_t+k+1} - u_{(n-1)n_t+k} \right]^2 + \frac{1}{n_t} \sum_{k=1}^{n_t} \theta_k u_{(n-1)n_t+k}^2 \right\}. \tag{3.2}$$

Here we set $\theta_k := \theta(t_k)$ for the user-defined weighting coefficients.

To solve the resulting constrained nonconvex optimization problem, the interior point based software package IPOPT [32] was chosen. We developed a C++ interface that calls the Sundance state and adjoint solvers to compute the cost function $j(u)$ and its derivative $j'(u)$. The interior point algorithm in IPOPT is configured to use limited-memory BFGS updates [36] for approximating the Hessian matrix of the Lagrange function.

4. Numerical methods for the state equation. Denoting by y^k the time-discretized state variables at time $t = t_k$ and applying an implicit Euler scheme to the state system (2.2), we obtain the following semi-discretized state equation in strong form:

$$\phi \frac{b^i(y^k) - b^i(y^{k-1})}{\Delta t} - \operatorname{div} B^i(y^k, \nabla y^k, x) - q^i\big|_{t=t_k} = 0 \quad (i = 1,2) \tag{4.1}$$

together with the algebraic complementarity relation $g(y^k) = 0$ and the boundary conditions for $1 \le k \le n_t$; the initial condition $y^0 = y_d$ reproduces the Dirichlet values.

Concerning the space discretization, we use the BOX method [2], a control-volume FE method. It has two main advantages: On the one hand it is locally conservative due to its control-volume nature, on the other hand it possesses sufficient stability due to involved full upwinding [15] and mass lumping [18]. Both the control-volume and upwinding features were integrated into the PDE toolbox Sundance [22] within the Trilinos framework to make it accessible for such stabilized locally conservative methods.

The BOX method uses piecewise linear continuous finite elements on a primal (in our case triangular) mesh for the discretization of the state y^k. Similar to finite volume methods, constant functions on the vertex-centered cells ("boxes") of the dual mesh are used as test functions; see Figure 4.1 for the construction of a box in the dual mesh.

Given a primal triangulation of the domain $\Omega \subset \mathbb{R}^2$, we denote the collection of boxes in the dual mesh by B_j ($j = 1, \dots, m$) and their respective box volumes by ω_j. To derive the space discretization, the semi-discrete state equation (4.1) is integrated over each box B_j – note that the indicator functions $\mathbf{1}_{B_j}$ are a basis of our test space – and Gauß' divergence theorem is applied to receive a boundary integral:

$$\begin{aligned} &\int_{B_j} \left[\phi \frac{b^i(y^k) - b^i(y^{k-1})}{\Delta t} - \operatorname{div} B^i(y^k, \nabla y^k, x) - q^i\big|_{t=t_k} \right] dx \\ &= \int_{B_j} \left[\phi \frac{b^i(y^k) - b^i(y^{k-1})}{\Delta t} - q^i\big|_{t=t_k} \right] dx - \int_{\partial B_j} n^T B^i(y^k, \nabla y^k, x)\, dS = 0 \end{aligned} \tag{4.2}$$

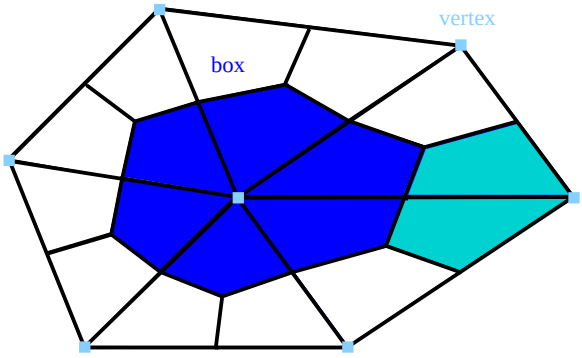

FIG. 4.1. *Construction of the dual mesh.*

Inserting piecewise linear Lagrange finite elements for the state and applying upwinding to the term $n^T B^i$ in the boxwise weak form (4.2), we then arrive at

$$\int_{B_j} \left[\phi \frac{b^i(y^k) - b^i(y^{k-1})}{\Delta t} - q^i \big|_{t=t_k} \right] dx - \int_{\partial B_j} n^T B^i_{\mathrm{up}} \, dS = 0 \quad (i = 1, 2). \tag{4.3}$$

The upwinding procedure for the modified terms B^i_{up} in the boundary integrals of (4.3) will be explained below. Quadrature rules for the boxes are chosen as follows: Volume integrals are evaluated vertex-centered, i.e.,

$$\int_{B_j} f(x) \, dx \approx \omega_j f(v_j), \quad 1 \le j \le m,$$

where v_j denotes the unique primal vertex inside the box B_j (see Figure 4.1). For the underlying control-volume FE method this yields a mass lumping, where the entries of the FE mass matrix are assigned to its main diagonal [18]. The use of the lumped mass matrix counteracts the appearance of non-physical oscillations in the solution [10].

The boundary integrals $\int_{\partial B_j} n^T B^i_{\mathrm{up}} \, dS$ are evaluated via a midpoint rule on each line segment of the polygonal curve ∂B_j in triangular 2D grids. The fully implicit upwinding procedure is based on the decomposition

$$B^i = -\sum_\alpha C^i_\alpha V_\alpha + D^i \rho_w X^i_w \quad \text{with} \quad C^i_\alpha := \lambda_\alpha \rho_\alpha X^i_\alpha, \quad V_\alpha := -\mathbf{K}\big(\nabla p_\alpha + \rho_\alpha \mathfrak{g} e_2\big).$$

Upon implicit evaluation of the Darcy velocities V_α at the discretized state y^k, the above advective coefficients C^i_α are implicitly evaluated

- at the vertex v_j of the box B_j for outflow $n^T V_\alpha \ge 0$,
- at the vertex of the opposite box for inflow $n^T V_\alpha < 0$.

Here n denotes the outer dual cell normal of the box B_j and "opposite" in 2D is subject to the respective line segment of the polygonal curve ∂B_j (see Figure 4.1). The evaluation of the upwinded flux terms B^i_{up} is based on this strategy. For further details we refer the reader to Chapter 3 in [2] for two-phase flow or Section 4.2 in [4] for miscible flow.

5. Implementation strategies in Trilinos/Sundance. As a software platform for state and adjoint simulations, the choice was made for the FE library package Sundance [22], based

on the scientific computing framework Trilinos [16]. Within the MAC–TUM project B7, Sandia National Labs, Texas Tech University and Technische Universität München (TUM) have joined forces to further develop and improve this toolbox, especially in view of PDE-constrained optimization and multiphysics problems [3, 23]. The mentioned C++ interface of the Sundance simulation code to IPOPT is based on earlier work on shape optimization with Navier–Stokes flow [5, 6, 21].

Sundance provides a wide range of benefits, such as efficient treatment of nonlinearities, automatic generation of linearized equations and thereby support for implementing adjoints within efficient gradient based optimization algorithms [23]. In our context, we represent the whole optimal control problem by its Lagrange function $\mathcal{L}(y, u, \mu)$ as a timestep-wise variational form in Sundance. Here μ denotes the adjoint state, i.e., the Lagrange multiplier of the state equation. Sundance conveniently provides routines that automatically derive nonlinear and linear variational problems from this corresponding to $\mathcal{L}_{\mu^k} = 0$, which is the state equation, and $\mathcal{L}_{y^k} = 0$, which is the adjoint equation. Furthermore, Sundance is MPI based and provides parallel system assembly as well as interfaces to a whole library of direct and iterative parallel solvers (e.g., NOX, Amesos, AztecOO).

In order to make these convenient features available for our purposes, we had to implement several additional class structures and quadrature rules within the PDE toolbox, since Sundance so far had not directly supported control volumes and upwinding, as explained in Section 4. Since those two issues are essential ingredients of the BOX method, we integrated a 2D implementation of these features into the object oriented C++ framework of Sundance; an extension to 3D can be achieved with moderate efforts.

In the following, we illustrate the mentioned advantages for the implementation of state and adjoint equation systems: The state equation for y^k and its adjoint for μ^k at time step t_k can be written in terms of the time-discretized Lagrangian:

$$\begin{aligned}
&\mathcal{L}_{\mu^k}(y, u, \mu) = 0 \quad \text{(time-discretized state equation)},\\
&\mathcal{L}_{y^k}(y, u, \mu) = 0 \quad \text{(corresponding adjoint equation)},\\
&\text{where} \quad y := \big(y^1, \dots, y^{n_t}\big), \quad \mu := \big(\mu^1, \dots, \mu^{n_t}\big).
\end{aligned}$$

This Lagrangian $\mathcal{L}$ contains the stepwise weak state equation functionals $\langle \mu^k, E_k \rangle$ and the stepwise evaluations of the objective functional (2.5) as follows:

$$\mathcal{L}(y, u, \mu) = \sum_{k=1}^{n_t} L_k\big(y^k, y^{k-1}, u, \mu^k\big) \quad \text{with timestep-wise Lagrangian}$$

$$L_k := \Delta t \, \langle \mu^k, E_k\big(y^k, y^{k-1}, u\big) \rangle + \delta_{kn_t} \int_\Omega f\big(y^{n_t}\big)\, dx - \gamma r_k(u),$$

where f is an abbreviation for the final-time density function in the integral of (2.5) and the terms r_k denote the stepwise components of the time-discretized penalty term (3.2).

Since μ^k only appears in the functional L_k, whereas y^k appears both in L_k and L_{k+1} for $1 \le k < n_t$ and only in L_{n_t} for $k = n_t$, this means that, defining two further functionals F and G at a respective fixed time step t_k via

$$F := L_k \quad \text{and} \quad G := L_k + (1 - \delta_{kn_t}) L_{k+1},$$

one can write the weak time-discretized state and adjoint equations as $F_{\mu^k} = 0$ and $G_{y^k} = 0$. This provides the basis for our flexible implementation of the discrete state system on the one hand and its exact discrete adjoint system on the other hand.

Sundance allows to implement variational forms in a very concise way. For example,

```
Expr auv = Integral(Omega, (grad*u)*(grad*v), quad);
```

generates the bilinear form corresponding to $a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx$, using the quadrature rule `quad`. Here, `u` and `v` are FE functions, e.g., given by

```
Expr u = new UnknownFunction(new Lagrange(1), "u");
```

as piecewise linear Lagrange elements. Forms representing Dirichlet boundary conditions can be defined with the method `EssentialBC` instead of `Integral`. The command

```
Functional L(mesh, intform, bcform, vecType);
```

combines the `Integral` form `intform` and the `EssentialBC` form `bcform` to construct a functional `L`. Here `mesh` describes the computational grid and `vecType` is the Trilinos vector type. Along these lines the functionals F and G can be implemented in Sundance as `Functional F` and `G`, respectively. For efficiency reasons we include only those parts of L_{k+1} in G that depend on the current state y^k. The methods

```
NonlinearProblem F.nonlinearVariationalProb(mu,mus,y,ys,...);
LinearProblem G.linearVariationalProb(y,ys,mu,...);
```

generate a nonlinear problem corresponding to the state timestep equation $F_{\mu^k}(y^k) = 0$ and a respective linear problem for the adjoint timestep equation $G_{y^k}(\mu^k) = 0$. Here `y` is the unknown vector of the state variables y^k (with several additional components to facilitate upwinding) and `mu` denotes a corresponding adjoint variable vector. The FE functions `ys` and `mus` provide storage for `y` and `mu`, while the dots hide further fixed parameters in the respective equation, for instance y^{k+1} and μ^{k+1} in $G_{y^k}(\mu^k) = 0$.

The order of parameters in `nonlinearVariationalProb` is: variable w.r.t. which the functional shall be linearized (here: `mu`); value of the linearization point (here: `mus`); unknown variable for which the resulting equation shall be solved (here: `y`); storage for previous variable (here: `ys`); fixed variables (here: not written out); values of the fixed variables (here: not written out). The order of parameters for `linearVariationalProb` applies accordingly, but no storage is provided for the unknown variable.

For runtime acceleration, we implemented an IMPES-like strategy to generate good initial iterates for implicit time steps that turn out to require too many NOX iterations (see [28] for details). Furthermore, the state and adjoint simulation was parallelized by domain decomposition. Tests were run on up to $n_p = 32$ processor units with good strong scaling properties even on relatively coarse grids. The parallelization uses in-built features of Sundance with MPI based communication between processors.

6. Discussion of optimization results. SI units are suppressed during the discussion. The following case study was performed in a rectangular reservoir $\Omega = [0, a_1] \times [0, a_2]$ of height $a_2 = 100$ and width $a_1 = 4a_2$, having the Neumann boundary $\Gamma_n = [0, a_1] \times \{0, a_2\}$ on the top and bottom and the Dirichlet boundary $\Gamma_d = \{0, a_1\} \times [0, a_2]$ on the side strips. Dirichlet conditions are chosen as $S_w = 1$, $p_n = 8 \cdot 10^6$ on top (otherwise hydrostatic) and $X \approx 10^{-4}$, so that complementarity $g(y) = 0$ holds on this boundary.

Moreover, we restrict our attention to a reservoir with constant porosity $\phi = 0.2$ and impose constant residual saturations $S_w^r = S_n^r = 0.1$ for both water and CO_2. The Sundance simulation is done on a triangular grid with $n_1 \times n_2$ uniform rectangles in Ω (two triangles

per rectangle). We stress, however, that our implementation supports unstructured grids. In what follows, a horizontal domain decomposition with $n_p = 16$ parallel MPI processes is applied on a mesh with $n_1 = 128$ and $n_2 = 32$.

The rectangular model reservoir is supplied with a heterogeneous isotropic absolute permeability distribution $\mathbf{K}(x) = k_0(x)\mathbf{id}$ and with $N = 5$ injection wells. The function $k_0(x)$ is illustrated in Figure 6.1: The permeability increases from left to right and has a horizontal layer of significantly lowered permeability. The five wells are ordered in a trapezoidal pattern. Notice that the two upper wells are located within the blue layer of low permeability.

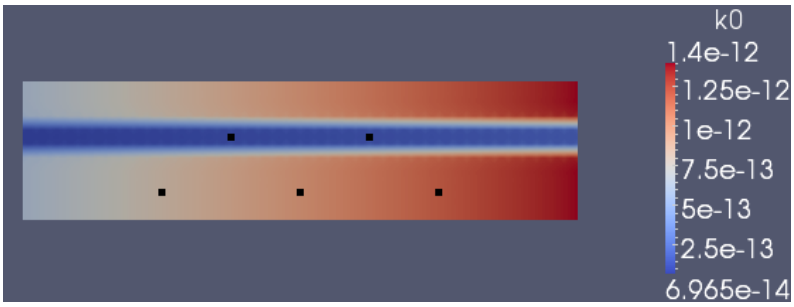

FIG. 6.1. *Permeability distribution k_0 and injection well locations.*

The optimization was run for $T = 4$ months with $n_t = 200$ time steps and has converged up to an IPOPT tolerance of $2 \cdot 10^{-5}$, in our case corresponding to a dual infeasibility norm (see [32] for precise definitions); the objective function was scaled by -10^{-3} in the interface, since IPOPT minimizes and expects problem-adjusted scaling. We set $\beta_1 = \beta_2 = 1$ in (2.5) for a combination of residual and solubility trapping in the objective, while the regularization parameter is $\gamma = 5 \cdot 10^{-4}$ and the weight function $\theta(t) = 4 \cdot 10^3\big(1 + \frac{4t}{T}\big)$ in (2.6) penalizes late injection. Given a scaling of $q_* = 200$ in (3.1), the control constraints

$$0 \le u_j \le 9 \quad \text{and} \quad c(u) = \frac{1}{n_t} \sum_{j=1}^{Nn_t} u_j \le 18$$

make sure that at least two wells must be active for maximal injection $c(u) = 18$.

Figures 6.2–6.4 illustrate the final saturation profiles for the computed optimal injection strategy in comparison to an early bang-bang strategy – i.e., inject the maximal amount in all wells for the first 80 time steps, then switch injection off – and to uniform injection with $c(u) = 18$. In contrast to other results in [28], the optimal control has $c(u) \approx 17.856 < 18$, so the upper bound is not active in this situation. This behavior shows that one must not inject too much CO_2 into the reservoir. The objective functional values are:

- optimal control: $J = 2.465 \cdot 10^5$, $J\big|_{\gamma=0} = 2.471 \cdot 10^5$,
- early bang-bang strategy: $J\big|_{\gamma=0} = 2.429 \cdot 10^5$,
- uniform injection ($u_j = 3.6$): $J\big|_{\gamma=0} = 1.965 \cdot 10^5$.

This shows that the early bang-bang strategy comes quite close to the optimal solution and is significantly superior to uniform injection. Accordingly, Figures 6.2 and 6.3 look very similar, except for the red brine hole below the upper CO_2 plume in Figure 6.2. Note that, in contrast to the uniform strategy in Figure 6.4, the three lower "pillars" of nonwetting supercritical CO_2 in Figures 6.2 and 6.3 are cut off from the upper portion of the supercritical plume, being immobilized below the residual saturation $S_n^r = 0.1$.

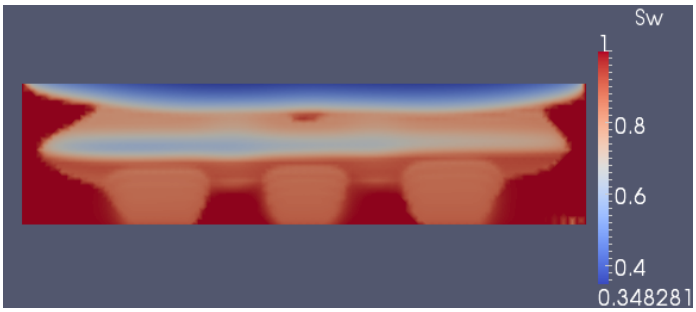

FIG. 6.2. *Final saturation S_w for optimal control.*

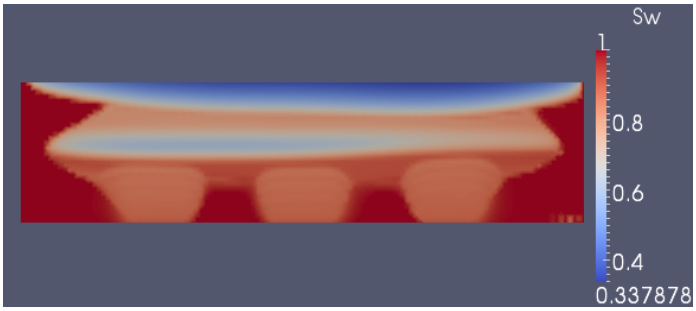

FIG. 6.3. *Final saturation S_w for early bang-bang.*

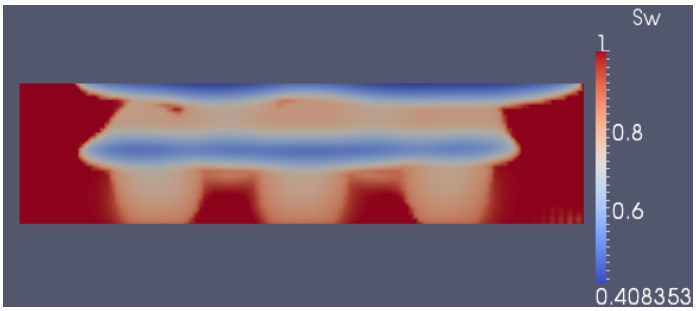

FIG. 6.4. *Final saturation S_w with uniform $u_j = 3.6$.*

The optimal injection rates in the upper and lower wells are shown in the respective Figures 6.5 and 6.6. We recognize the following tendencies: The rates come rather close to "on–off" bang-bang strategies, while the upper wells switch back to maximal injection shortly before the end. This can be attributed to the fact that they are located in a layer of low permeability, so parts of the final portion of injected CO_2 can remain below residual saturation, due to a slower CO_2 plume evolution.

The mostly bang-bang like structure – i.e., either maximal or minimal injection in every well at all times ("on–off") – of the optimal strategies (with moderate deviation for the lower middle well) is supported by the recent article [34]: There it is shown that, given optimal control of EOR water-flooding with immiscible two-phase flow, such bang-bang solutions may occur if the control constraints on the injection rates are linear. Further optimization results for miscible two-phase flow in [28] agree with this observation.

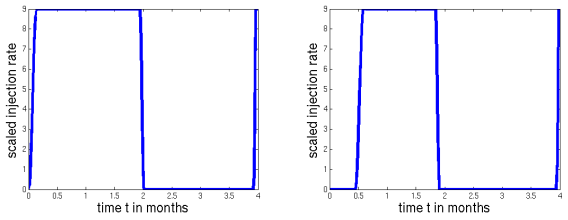

FIG. 6.5. *Optimal rates in the upper two wells.*

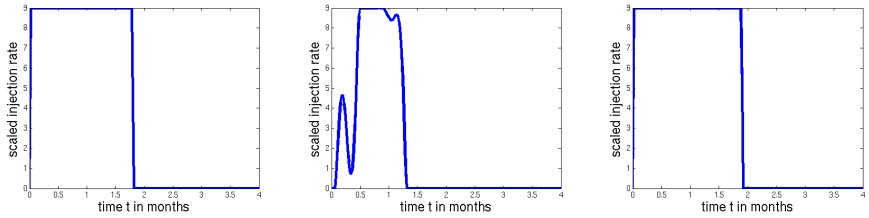

FIG. 6.6. *Optimal rates in the lower three wells.*

Acknowledgements. This publication is based on work supported by Award No. UK-C0020, made by King Abdullah University of Science and Technology (KAUST). The work was conducted for the MAC–KAUST project K1 "Simulating CO_2 Sequestration" within the Munich Centre of Advanced Computing (MAC) at TUM. In addition, the authors gratefully acknowledge support by DFG INST 95/919-1 FUGG.

REFERENCES

[1] H. Asheim: *Maximization of water sweep efficiency by controlling production and injection rates*; SPE paper 18365, presented at "SPE European Petroleum Conference", London, 1988.

[2] P. Bastian: *Numerical Computation of Multiphase Flows in Porous Media*; Habilitation thesis, Christian–Albrechts-Universität Kiel, 1999.

[3] J. Benk, M. Mehl, M. Ulbrich: *Sundance PDE solvers on Cartesian fixed grids in complex and variable geometries*; in Proceedings of "ECCOMAS Thematic Conference on CFD and Optimization", Antalya, 2011.

[4] A. Bielinski: *Numerical Simulation of CO_2 Sequestration in Geological Formations*; Dissertation, Universität Stuttgart, 2006.

[5] C. Brandenburg, F. Lindemann, M. Ulbrich, S. Ulbrich: *A continuous adjoint approach to shape optimization for Navier Stokes flow*; in K. Kunisch et al. (eds.), Optimal Control of Coupled Systems of Partial Differential Equations, Int. Ser. Numer. Math. 158, Birkhäuser, Basel, 35–56, 2009.

[6] C. Brandenburg, F. Lindemann, M. Ulbrich, S. Ulbrich: *Advanced numerical methods for PDE constrained optimization with application to optimal design in Navier Stokes flow*; in S. Engell et al. (eds.), Constrained Optimization and Optimal Control for Partial Differential Equations, Birkhäuser, Basel, 257–275, 2011.

[7] R.H. Brooks, A.T. Corey: *Hydraulic properties of porous media*; Hydrology Papers 3, Colorado State University, Fort Collins, 1964.

[8] D.R. Brouwer: *Dynamic Water Flood Optimization with Smart Wells Using Optimal Control Theory*; PhD thesis, Delft University of Technology, 2004.

[9] D.R. Brouwer, J.D. Jansen: *Dynamic optimization of water flooding with smart wells using optimal control theory*; SPE J. 9(4), 391–402, 2004.

[10] M.A. Celia, P. Binning: *A mass-conservative numerical solution for two-phase flow in porous media with application to unsaturated flow*; Water Resour. Res. 28(10), 2819–2828, 1992.

[11] Z. Chen, G. Huan, Y. Ma: *Computational Methods for Multiphase Flows in Porous Media*; Computational Science and Engineering Series, vol. 2, SIAM, 2006.

[12] N. Dolle, D.R. Brouwer, J.D. Jansen: *Dynamical optimization of water flooding with multiple injectors and producers using optimal control theory*; in Proceedings of "14th International Conference on Computational Methods in Water Resources", Delft, 2002.
[13] A. Fischer: *Solution of monotone complementarity problems with locally Lipschitzian functions*; Math. Program. 76(3B), 513–532, 1997.
[14] G. Gao, A.C. Reynolds: *An improved implementation of the LBFGS algorithm for automatic history matching*; SPE J. 11(1), 5–17, 2006.
[15] R. Helmig: *Multiphase Flow and Transport Processes in the Subsurface*; Springer, 1997.
[16] M.A. Heroux, J.M. Willenbring, R. Heaphy: *Trilinos developers guide*; Sandia National Laboratories, SAND 2003–1898, 2003.
[17] R. Helmig, J. Niessner, H. Class: *Recent advances in finite element methods for multiphase flow processes in porous media*, Int. J. Comput. Fluid Dyn. 20(3), 245–252, 2006.
[18] R. Huber, R. Helmig: *Multiphase flow in heterogeneous porous media: A classical finite element method versus an implicit pressure–explicit saturation-based mixed finite element–finite volume approach*; Int. J. Numer. Methods Fluids 29(8), 899–920, 1999.
[19] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich: *Optimization with PDE Constraints*; Mathematical Modelling: Theory and Applications, vol. 23, Springer, 2008.
[20] J.D. Jansen: *Adjoint-based optimization of multi-phase flow through porous media – a review*; Comput. Fluids 46(1), 40–51, 2011.
[21] F. Lindemann: *Theoretical and Numerical Aspects of Shape Optimization with Navier-Stokes Flows*; Dissertation under review, Technische Universität München, 2012.
[22] K.R. Long: *Sundance rapid prototyping tool for parallel PDE optimization*; in L.T. Biegler et al. (eds.), Large scale PDE-constrained optimization, Springer, 331–342, 2003.
[23] K.R. Long, P.T. Boggs, B.G. van Bloemen Wanders: *Sundance: High-level software for PDE-comstrained optimization*; SAND report, preprint, 2012.
[24] G.J. Mehos, W.F. Ramirez: *Use of optimal control theory to optimize carbon dioxide miscible-flooding enhanced oil recovery*; J. Petroleum Sci. Eng. 2(4), 247–260, 1989.
[25] D.S. Oliver, A.C. Reynolds, N. Liu: *Inverse Theory for Petroleum Reservoir Characterization and History Matching*; Cambridge University Press, 2008.
[26] M. Peszynska, E.W. Jenkins, M.F. Wheeler: *Boundary conditions for fully implicit two-phase flow models*; in X. Feng, T.P. Schulze (eds.), Recent Advances in Numerical Methods for Partial Differential Equations and Applications, Contemporary Mathematics Series 306, AMS, 85–106, 2002.
[27] P. Sarma, K. Aziz, L.J. Durlofsky: *Implementation of adjoint solution for optimal control of smart wells*; SPE paper 92864, presented at "SPE Reservoir Simulation Symposium", Houston, 2005.
[28] M. Simon, M. Ulbrich: *Adjoint based optimal control of partially miscible two-phase flow in porous media with applications to CO_2 sequestration in underground reservoirs*; preprint in preparation, Technische Universität München, to appear, 2012.
[29] B. Sudaryanto, Y.C. Yortsos: *Optimization of fluid front dynamics in porous media using rate control*; Phys. Fluids 12(7), 1656–1670, 2000.
[30] B. Sudaryanto, Y.C. Yortsos: *Optimization of displacements in porous media using rate control*; SPE paper 71509, presented at "SPE Annual Technical Conference and Exhibition", New Orleans, 2001.
[31] G.A. Virnovski: *Water flooding strategy design using optimal control theory*; in Proceedings of "6th European Symposium on IOR", Stavanger, 1991.
[32] A. Wächter, L.T. Biegler: *On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming*; Math. Program. 106, 25–57, 2006.
[33] I.S. Zakirov, S.I. Aanonsen, E.S. Zakirov, B.M. Palatnik: *Optimization of reservoir performance by automatic allocation of well rates*; in Proceedings of "ECMOR V: European Conference on Mathematics of Oil Recovery", Leoben, 1996.
[34] M.J. Zandvliet, O.H. Bosgra, J.D. Jansen, P.M.J. van den Hof, J.F.B.M. Kraaijewanger: *Bang-bang control and singular arcs in reservoir flooding*; J. Petroleum Sci. Eng. 58(1), 186–200, 2007.
[35] Q. Zhang: *A finite difference–streamline diffusion (FDSD) method and its error estimates for two-phase incompressible miscible flow in porous media*; Acta Math. Appl. Sin. 26(2), 318–327, 2003.
[36] C. Zhu, R.H. Byrd, J. Nocedal: *L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization*, ACM Transactions on Mathematical Software 23(4), 550–560, 1997.