THE FEASIBILITY OF A NEW OPTIMIZATION APPROACH TO MULTIPHASE FLOW *

LAWRENCE C. COWSAR †, ROLAND GLOWINSKI‡, ANTHONY J. KEARSLEY§, MARY F. WHEELER¶ AND IVAN YOTOV ‖

Abstract. A new optimization formulation for multiphase flow in porous media is introduced. A locally mass conservative mixed finite element method is used for the spatial discretization. An unconditionally stable, fully implicit time discretization is also used and leads to a coupled system of nonlinear equations that must be solved at each time step. We reformulate this system as a least squares problem with simple bounds involving only one of the phase saturations. Both a Gauss-Newton method and a BFGS secant method are considered as potential solvers for the optimization problem. Each evaluation of the least squares objective function and gradient requires solving two single-phase self-adjoint, linear, uniformly elliptic partial differential equations for which very efficient solution techniques have been developed. However, numerical experiments suggest that the new formulation may not be computationally competitive with existing methods.

Key Words. Multiphase flow, mixed methods, cell-centered finite differences

1. Introduction

The flow of two immiscible fluids is modeled by a system of nonlinear transient partial differential equations coupled with nonlinear algebraic constitutive relationships. Spatial discretization combined with stable fully implicit time stepping leads to a nonlinear system of algebraic equations that must be solved at every time step. Even though the differential operators involved are self-adjoint, the use of Newton-type solution methods require the solution of a system of nonsymmetric linear equations to calculate each Newton step. The objective of this work is to develop an algorithm that exploits the fact that the operators are self-adjoint and only requires the solution of linear systems that are symmetric and positive definite. Our work is predicated on the observation that, all other things being equal, it is easier to solve symmetric positive definite linear systems than more general nonsymmetric systems. Unfortunately, we will find that in the new formulation all other things are not equal.

In this work we consider immiscible two-phase flow in a porous medium. Let $\Omega \subset \mathbf{R}^d$, $d = 2, 3$ be the domain of the problem. The equations governing fluid motion through porous medium are mass conservation of the phases [4, 7],

$$\frac{\partial(\phi s_i \rho_i)}{\partial t} + \nabla \cdot \rho_i \mathbf{u}_i = q_i, \quad (x,t) \in \Omega \times (0,T), \tag{1.1}$$

where $i = w$ (wetting), n (non-wetting) denotes the phase, s_i is the phase saturation, ρ_i is the

*This work was supported in part by the Department of Energy, the National Science Foundation, and the State of Texas Governor's Energy Office. This document is in its final form and will not appear elsewhere.

†AT&T Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974.

‡Department of Mathematics, University of Houston, Houston, Texas 77204.

§Department of Mathematics, University of Massachusetts at Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747.

¶Texas Institute of Computational and Applied Mathematics, University of Texas at Austin, Austin, TX 78712.

‖Department of Computational and Applied Mathematics, Rice University, 6100 Main Street, Houston, TX 77005, and Texas Institute of Computational and Applied Mathematics, University of Texas at Austin, Austin, TX 78712.

phase density, ϕ is the porosity, q_i is the source term, and

$$\mathbf{u}_i = -\frac{k_i(s_i)K}{\mu_i}(\nabla p_i - \rho_i g \nabla D), \quad (x,t) \in \Omega \times (0,T), \tag{1.2}$$

is the Darcy velocity. Here p_i is the phase pressure, K is the absolute permeability tensor, $k_i(s_i)$ is the phase relative permeability, μ_i is the phase viscosity, g is the gravitational constant, and D is the depth. The two equations are coupled with the volume balance equation

$$s_w + s_n = 1, \tag{1.3}$$

and the capillary pressure relation

$$p_c(s_w) = p_n - p_w. \tag{1.4}$$

No flow boundary conditions

$$\mathbf{u}_i \cdot \nu = 0, \quad (x,t) \in \partial\Omega \times (0,T), \tag{1.5}$$

where ν is the outward pointing unit normal vector on $\partial\Omega$, and initial condition

$$s_w(x,0) = s_0(x), \quad x \in \Omega, \tag{1.6}$$

complete the model. We note that the no flow boundary conditions are assumed only for simplicity of the presentation. More general boundary conditions can be treated with the same techniques. Gravity effects will also be neglected for the sake of simplicity. We also assume that the fluids and the media are incompressible, i.e. ρ_i and ϕ are constants.

The main idea in our approach is to decouple the two conservation equations. We notice that the only coupling is through the relative permeabilities $k_i(s_i)$. If the saturation of one of the phases is known, the equations are decoupled. Moreover, considered as equations for phase pressures, they are linear elliptic equations that can be solved easily. This observation motivates the optimization formulation below.

REMARK 1.1. Throughout the paper we refer to the pressure equations as elliptic equations. In fact, they are degenerate elliptic, since the relative permeability $k_i(s_i) = 0$ for $s_i = 0$, $i = w, n$. In order to obtain a symmetric positive definite matrix after the discretization, we modify the equations as explained in Section 3.

REMARK 1.2. Even when the elliptic coefficients are strictly positive, the resulting pressure matrices are positive semidefinite because of the no flow boundary conditions. We assume that there is at least one well with specified bottom hole pressure which leads to nonsingular matrices and uniquely determines the pressures.

Let $s = s_w$ and define the nonlinear operator F by

$$F(s) = p_c(s) - (p_n(s) - p_w(s)).$$

Here $p_n(s)$ and $p_w(s)$ are the solutions to the linear elliptic equations (1.1)–(1.2) with (1.3) and (1.5) for the given s. We have suppressed the time dependence of F inherited from the time dependence of s and the dependence of $p_n(s)$ and $p_w(s)$ on $\partial s/\partial t$. From F we derive a time dependent nonlinear functional f by setting

$$f(s) = \frac{1}{2}\|F(s)\|^2_{L^2(\Omega)} \equiv \frac{1}{2}\int_\Omega |F(s(x))|^2 dx \tag{1.7}$$

and consider the minimization problem

$$\min_{s \in L^2(\Omega)} f(s).$$

By construction, the minimum of $f(s)$ is identically zero for almost every time and occurs at the solution to (1.1)–(1.5).

The rest of the paper is organized as follows. The mixed finite element discretization is described in Section 2. In Section 3 we present the optimization formulation of the discretized problem. The optimization methods employed are described in Section 4. We close with some numerical experiments and conclusions in Section 5.

2. Discretization

In this section we present the expanded mixed finite element method used for the spatial discretization of the multiphase flow system combined with a backward Euler temporal discretization. The expanded mixed method (see [22, 16, 8, 2]) is desirable because it conserves mass locally (on every element), provides accurate velocities, and allows for efficient implementation even when the permeability is a possibly degenerate full tensor. Quadrature rules are used for the numerical integration leading to cell-centered finite difference equations for the pressures and saturations. The relationship between cell-centered finite differences and mixed methods were first noticed by Russell and Wheeler [20] and extends to the expanded mixed method [2].

The backward Euler method is used to discretize the time derivative resulting in a fully implicit time stepping scheme. The fully implicit discretization is used because all explicit or partially implicit schemes suffer from severe time step stability limitations when advection processes are dominating. One point upstream weighting is used to further stabilize the method.

2.1. An expanded mixed finite element method

Let us define

$$H(\mathrm{div};\Omega) = \{\mathbf{v} \in (L^2(\Omega))^d : \nabla \cdot \mathbf{v} \in L^2(\Omega)\},$$

with the norm

$$\|\mathbf{v}\|_{H(\mathrm{div};\Omega)} = \left\{ \int_\Omega \left(|\mathbf{v}|^2 + |\nabla \cdot \mathbf{v}|^2\right) dx \right\}^{1/2},$$

and set

$$\mathbf{V} = \{\mathbf{v} \in H(\mathrm{div};\Omega) : \mathbf{v} \cdot \nu = 0 \text{ on } \partial\Omega\}.$$

Following [2], we introduce the pressure gradients

$$\tilde{\mathbf{u}}_i = -\nabla p_i, \quad i = w, n.$$

Then

$$\mathbf{u}_i = \lambda_i(s) K \tilde{\mathbf{u}}_i,$$

where $\lambda_i = k_i/\mu_i$ is the phase mobility.

For each time $t \in (0,T]$, in the expanded mixed variational form we seek $\mathbf{u}_i(\cdot,t) \in \mathbf{V}$, $\tilde{\mathbf{u}}_i(\cdot,t) \in (L^2(\Omega))^d$, $p_i(\cdot,t) \in L^2(\Omega)$, $i = w,n$, and $s(\cdot,t) \in L^2(\Omega)$ such that

$$(\mathbf{u}_i, \tilde{\mathbf{v}}) = (\lambda_i(s)K\tilde{\mathbf{u}}_i, \tilde{\mathbf{v}}), \qquad \tilde{\mathbf{v}} \in (L^2(\Omega))^d, \tag{2.1}$$

$$(\tilde{\mathbf{u}}_i, \mathbf{v}) - (p_i, \nabla\cdot\mathbf{v}) = 0, \qquad \mathbf{v} \in \mathbf{V}, \tag{2.2}$$

$$\phi(\frac{\partial s}{\partial t}, w) + (\nabla\cdot\mathbf{u}_w, w) = (\tilde{q}_w(s), w), \qquad w \in L^2(\Omega), \tag{2.3}$$

$$-\phi(\frac{\partial s}{\partial t}, w) + (\nabla\cdot\mathbf{u}_n, w) = (\tilde{q}_n(s), w), \qquad w \in L^2(\Omega), \tag{2.4}$$

$$(p_n - p_w, w) = (p_c(s), w), \qquad w \in L^2(\Omega), \tag{2.5}$$

$$(s(\cdot,0), w) = (s_0, w), \qquad w \in L^2(\Omega), \tag{2.6}$$

where $\tilde{q}_i = q_i/\rho_i$ and $(\,\cdot\,,\,\cdot\,)$ denotes the $L^2(\Omega)$ inner product.

Let $\mathcal{T}_h$ be a finite element partition of Ω where h is associated with the size of the elements. Let $\mathbf{V}_h \times W_h \subset \mathbf{V} \times L^2(\Omega)$ be any of the known mixed finite element spaces on $\mathcal{T}_h$ (see [5] for a detailed description of these spaces). Let $\tilde{\mathbf{V}}_h$ be a finite element subspace of $(L^2(\Omega))^d$ such that $\mathbf{V}_h \subseteq \tilde{\mathbf{V}}_h$. Let $\{t_n\}_{n=0}^N$ be a partition of $[0,T]$ with $t_0 = 0$ and $t_N = T$, let $\Delta t^n = t_n - t_{n-1}$, and let $f^n = f(t_n)$.

We have the following backward Euler expanded mixed finite element method for the multiphase flow system. Find, for any $0 \le n \le N$, $(\mathbf{u}^n_{i,h}, \tilde{\mathbf{u}}^n_{i,h}, p^n_{i,h}) \in \mathbf{V}_h \times \tilde{\mathbf{V}}_h \times W_h$, and $s^n_h \in W_h$ such that

$$(\mathbf{u}^n_{i,h}, \tilde{\mathbf{v}}) = (\lambda_i(s^n_h)K\tilde{\mathbf{u}}^n_{i,h}, \tilde{\mathbf{v}}), \qquad \tilde{\mathbf{v}} \in \tilde{\mathbf{V}}_h, \tag{2.7}$$

$$(\tilde{\mathbf{u}}^n_{i,h}, \mathbf{v}) - (p^n_{i,h}, \nabla\cdot\mathbf{v}) = 0, \qquad \mathbf{v} \in \mathbf{V}_h, \tag{2.8}$$

$$\phi(\frac{s^n_h - s^{n-1}_h}{\Delta t^n}, w) + (\nabla\cdot\mathbf{u}^n_{w,h}, w) = (\tilde{q}_w(s^n_h), w), \qquad w \in W_h, \tag{2.9}$$

$$-\phi(\frac{s^n_h - s^{n-1}_h}{\Delta t^n}, w) + (\nabla\cdot\mathbf{u}^n_{n,h}, w) = (\tilde{q}_n(s^n_h), w), \qquad w \in W_h, \tag{2.10}$$

$$(p^n_{n,h} - p^n_{w,h}, w) = (p_c(s^n_h), w), \qquad w \in W_h, \tag{2.11}$$

$$(s^0_h, w) = (s_0, w), \qquad w \in W_h. \tag{2.12}$$

2.2. A cell-centered finite difference scheme

To obtain a finite difference method, we consider the lowest order Raviart-Thomas spaces [19] and take $\tilde{\mathbf{V}}_h = \mathbf{V}_h$ on the interior of Ω. We restrict our discussion to rectangular elements, noting that these techniques can be extended to logically rectangular and triangular grids (see [1]). Let

$$\tilde{\mathbf{V}}_h = \Big\{\mathbf{v} \in H(\mathrm{div};\Omega) : \mathbf{v}|_E = (\alpha_1 x_1 + \beta_1, \alpha_2 x_2 + \beta_2, \alpha_3 x_3 + \beta_3)^T, \alpha_i, \beta_i \in \mathbf{R}\Big\}$$

be the velocity space, and

$$W_h = \Big\{w \in L^2(\Omega) : w|_E = \alpha, \alpha \in \mathbf{R}\Big\}$$

be the pressure space. Let

$$\mathbf{V}_h = \Big\{\mathbf{v} \in \tilde{\mathbf{V}}_h : \mathbf{v}\cdot\nu = 0 \text{ on } \partial\Omega\Big\}.$$

We employ quadrature rules to approximate the vector integrals in (2.7) and (2.8). The two equations are replaced by

$$(\mathbf{u}^n_{i,h}, \tilde{\mathbf{v}})_T = (\lambda_i(s^n_h) K \tilde{\mathbf{u}}^n_{i,h}, \tilde{\mathbf{v}})_T, \qquad \tilde{\mathbf{v}} \in \tilde{\mathbf{V}}_h, \tag{2.13}$$

$$(\tilde{\mathbf{u}}^n_{i,h}, \mathbf{v})_T - (p^n_{i,h}, \nabla \cdot \mathbf{v}) = 0, \qquad \mathbf{v} \in \mathbf{V}_h, \tag{2.14}$$

where $(\cdot, \cdot)_T$ denotes an application of the trapezoidal rule to the $L^2(\Omega)$ inner product with respect to $\mathcal{T}_h$.

To avoid notation that is too complicated, we omit the phase index i and the time index n in (2.13) and (2.14) for the rest of this section. The application of the trapezoidal rule leads to diagonal coefficient matrices for $\mathbf{u}_h$ and $\tilde{\mathbf{u}}_h$ in (2.13) and (2.14). Therefore, $\mathbf{u}_h$ and $\tilde{\mathbf{u}}_h$ can be eliminated and a single equation for p_h can be obtained.

To describe the scheme we need some relatively standard cell-centered finite difference notation. We present the method for $d = 2$; a straightforward generalization exists for $d = 3$. Let

$$(x_{i+1/2}, y_{j+1/2}), \quad i = 0, ..., N_x, \; j = 0, ..., N_y,$$

be the grid points, and define, for $i = 1, ..., N_x$ and $j = 1, ..., N_y$,

$$x_i = \frac{1}{2}(x_{i+1/2} + x_{i-1/2}), \qquad y_j = \frac{1}{2}(y_{j+1/2} + y_{j-1/2}),$$

$$h^x_i = x_{i+1/2} - x_{i-1/2}, \qquad h^y_j = y_{j+1/2} - y_{j-1/2}.$$

We write $\mathbf{v} = (v^x, v^y)$ for $\mathbf{v} \in \mathbf{R}^2$, and for any function $g(x, y)$, let g_{ij} denote $g(x_i, y_j)$, let $g_{i+1/2,j}$ denote $g(x_{i+1/2}, y_j)$, etc.

Let us assume for simplicity that K is diagonal (see [2] for the full tensor case). Let $\mathcal{K}(s) = \lambda(s)K$. Choosing $\tilde{\mathbf{v}}$ in (2.13) to be the basis function at node $(i + 1/2, j)$ or at node $(i, j + 1/2)$, we get

$$\mathbf{u}^x_{h,i+1/2,j} = (\mathcal{K}_{11})_{i+1/2,j} \tilde{\mathbf{u}}^x_{h,i+1/2,j}, \qquad \mathbf{u}^y_{h,i,j+1/2} = (\mathcal{K}_{22})_{i,j+1/2} \tilde{\mathbf{u}}^y_{h,i,j+1/2}, \tag{2.15}$$

which is a finite difference approximation of $\mathbf{u} = \mathcal{K}\tilde{\mathbf{u}}$. The same choice of $\mathbf{v}$ in (2.14) gives

$$\tilde{\mathbf{u}}^x_{h,i+1/2,j} = -\frac{p_{h,i+1,j} - p_{h,i,j}}{\frac{1}{2}(h^x_i + h^x_{i+1})}, \qquad \tilde{\mathbf{u}}^y_{h,i,j+1/2} = -\frac{p_{h,i,j+1} - p_{h,i,j}}{\frac{1}{2}(h^y_j + h^y_{j+1})}, \tag{2.16}$$

which is a finite difference approximation of $\tilde{\mathbf{u}} = -\nabla p$. Finally, choosing w in (2.9) and (2.10) to be the basis function at element (i, j), we get

$$\left\{ \phi \frac{s^n_{h,i,j} - s^{n-1}_{h,i,j}}{\Delta t^n} + \frac{u^{x,n}_{w,h,i+1/2,j} - u^{x,n}_{w,h,i-1/2,j}}{h^x_i} + \frac{u^{y,n}_{w,h,i,j+1/2} - u^{y,n}_{w,h,i,j-1/2}}{h^y_j} \right\} h^x_i h^y_j$$

$$= \int_{E_{ij}} \tilde{q}_w(s^n_h) dx dy, \tag{2.17}$$

with a similar expression for (2.10).

The combination of (2.15), (2.16), and (2.17) produces the cell-centered finite difference approximation of $\phi \frac{\partial s_i}{\partial t} - \nabla \cdot \mathcal{K}_i(s_i) \nabla p_i = \tilde{q}_i$, $i = w, n$. The stencil for the pressures is 5 points for $d = 2$ and 7 points for $d = 3$ if K is a diagonal tensor. If K is a full tensor, the stencil is 9 points for $d = 2$ and 19 points for $d = 3$.

REMARK 2.1. Since $(\mathcal{K}_{11})_{i+1/2,j} = \lambda(s_h)_{i+1/2,j}(K_{11})_{i+1/2,j}$, equation (2.15) requires $\lambda(s_h)$ and K to be evaluated on the edges. Because s_h is constant on any element, $\lambda(s_h)_{i+1/2,j}$ cannot be directly computed. One-point upstream weighting is used to determine this value, i.e. $\lambda(s_h)_{i+1/2,j} = \lambda(s_h)_{i,j}$ or $\lambda(s_h)_{i+1/2,j} = \lambda(s_h)_{i+1,j}$ depending on the direction of the flow. If K is discontinuous and is given as a piecewise constant function, harmonic average is used for its value on the edges, i.e.

$$(K_{11})_{i+1/2,j} = \frac{2(K_{11})_{i,j}(K_{11})_{i+1,j}}{(K_{11})_{i,j} + (K_{11})_{i+1,j}}.$$

REMARK 2.2. The cell-centered finite difference scheme described above has been commonly used by the environmental and petroleum engineers [17]. It conserves the mass of all phases locally. Recognizing the relationship to the mixed finite element formulation provides more flexibility for handling irregular geometries and general boundary conditions. It is also useful in the convergence analysis.

3. A new optimization formulation

The cell-centered finite difference scheme presented in the previous section leads to a coupled system of nonlinear algebraic equations for one of the phase pressures and one of the saturations at every time step. In this section we propose a new optimization formulation of the discretized problem involving only the saturation.

Define a nonlinear operator $F : W_h \to W_h$ by

$$F(s_h^n) = p_c(s_h^n) - (p_n(s_h^n) - p_w(s_h^n)), \tag{3.1}$$

where $p_n(s_h^n)$ and $p_w(s_h^n)$ are the solution to the discrete system (2.13), (2.14), (2.9), (2.10) for a given s_h^n. Since we consider F in the context of implicit time stepping, we have suppressed the dependence of F on s_h^{n-1} and will drop the superscript n in what follows. Let $W(s_h) : W_h \to W_h$ be a symmetric linear operator depending on a function $s_h \in W_h$. We assume that $W(s_h)$ is uniformly positive definite with respect to all $s_h \in W_h$ and define a nonlinear nonnegative functional $f : W_h \to \mathbf{R}$ by

$$f(s_h) = \frac{1}{2}\int_\Omega (W(s_h)F(s_h))\, F(s_h)dx. \tag{3.2}$$

It is easy to see that solving the discrete system (2.9)–(2.11), (2.13), (2.14) for s_h^n is equivalent to finding s_h such that $f(s_h) = 0$ or to solving the minimization problem

$$\min_{s_h \in W_h} f(s_h). \tag{3.3}$$

REMARK 3.1. If the identity operator is used for the weighting operator $W(s)$, then (3.2) is the discrete analog of (1.7). We allow the possibility of a more general weighting function to improve the conditioning of the optimization problem. Several choices are given in Section 5.

3.1. Modification of the degenerate pressure equations

If only one phase is present in part of the domain, the pressure equation may be singular since the relative permeability for the phase vanishes at zero saturation. To overcome this obstacle, we introduce a modified system of equations. For a phase $i = n, w$, we write the discrete system (2.13), (2.14), (2.9), (2.10) in the form

$$A_i(S)P_i(S) = q_i(S), \tag{3.4}$$

where S is the vector of the saturation unknowns, P_i is the vector of pressures unknowns for phase i, $A_i(S)$ is the finite difference stiffness matrix, and $q_i(S)$ is the right hand side vector including the contribution from the time derivative of s.

Let $P_c(S) = p_c(S)$ denote the vector of capillary pressures calculated from the vector of saturations. Adding (3.4) for $i = n, w$ and using the fact that $P_c(S) = P_n(S) - P_w(S)$ at the solution due to (2.11), we see that

$$\left(A_n(S) + A_w(S)\right) P_n(S) = q_n(S) + q_w(S) + A_w(S)P_c(S).$$

Since the saturation of the wetting phase and the saturation of the non-wetting phase cannot both be zero in the same cell, it is easy to check that $(A_n(S) + A_w(S))$ is positive definite while $A_n(S)$ and $A_w(S)$ are only positive semi-definite in general. Adding equation (3.4) for $i = n$, we arrive at a nonsingular system that defines $P_n(S)$, namely,

$$\left(2A_n(S) + A_w(S)\right) P_n(S) = 2q_n(S) + q_w(S) + A_w(S)P_c(S). \tag{3.5}$$

A similar relationship holds for P_w; namely,

$$\left(A_n(S) + 2A_w(S)\right) P_w(S) = q_n(S) + 2q_w(S) - A_n(S)P_c(S). \tag{3.6}$$

By an abuse of notation, in what follows we will use A_i and q_i to denote the modified operators and right-hand sides. Since $p_c(s) = p_n - p_w$ at the solution, the modified system of equations has the same solution as the unmodified system in regions of nonzero saturation. In fact, the modified system simply guarantees that the pressure used for the absent phase is consistent with the the capillary pressure-saturation constitutive relation.

3.2. Evaluation of the objective function and its analytic gradient

The quasi-Newton methods employed for the solution of (3.3) are described in the next section. They require evaluation of the objective function and its gradient at each iteration. Given S, the evaluation of the objective function requires computing P_n and P_w by solving (3.5) and (3.6). Since the only nonlinearity and coupling of the two conservation equations is in $k_i(s_i)$, the equations are decoupled, linear, symmetric and, with the modification described above, positive definite.

To facilitate the exposition of the analytic gradient calculation, we write explicitly the dependence of the objective function f on P_n and P_w,

$$f(S) = f(S, P_n(S), P_w(S)). \tag{3.7}$$

By the chain rule we have

$$\frac{df}{dS} = \frac{\partial f}{\partial S} + \frac{\partial f}{\partial P_n}\frac{dP_n}{dS} + \frac{\partial f}{\partial P_w}\frac{dP_w}{dS}. \tag{3.8}$$

Differentiating (3.4), we find that $\frac{dP_i}{dS}$ satisfies

$$A_i\frac{dP_i}{dS} = \frac{dq_i}{dS} - \frac{dA_i}{dS}P_i, \tag{3.9}$$

an equation involving the same operator as the evaluation of the objective function. Combining (3.8) and (3.9) we get

$$\frac{df}{dS} = \frac{\partial f}{\partial S} + \frac{\partial f}{\partial P_n} A_n^{-1} \left(\frac{dq_n}{dS} - \frac{dA_n}{dS} P_n \right) + \frac{\partial f}{\partial P_w} A_w^{-1} \left(\frac{dq_w}{dS} - \frac{dA_w}{dS} P_w \right). \tag{3.10}$$

Therefore,

$$\begin{aligned} \nabla f = \left(\frac{\partial f}{\partial S} \right)^T + \left(\frac{dq_n}{dS} - \frac{dA_n}{dS} P_n \right)^T A_n^{-T} \left(\frac{\partial f}{\partial P_n} \right)^T \\ + \left(\frac{dq_w}{dS} - \frac{dA_w}{dS} P_w \right)^T A_w^{-T} \left(\frac{\partial f}{\partial P_w} \right)^T. \end{aligned} \tag{3.11}$$

Since the stiffness matrices A_i are symmetric, $A_i^{-T} = A_i^{-1}$. The partial derivative in (3.11) can be easily computed by the formulas

$$\frac{\partial f}{\partial S} = \frac{1}{2} F^T \left(\frac{dW}{dS} \right) F + (WF)^T \frac{dP_c}{dS}, \quad \frac{\partial f}{\partial P_n} = -(WF)^T, \quad \frac{\partial f}{\partial P_w} = (WF)^T.$$

The computation of the analytic gradient is dominated by the cost of solving two single phase, symmetric, linear, uniformly elliptic equations. Hence, the cost for evaluating ∇f is comparable to the one for evaluating f.

4. Solution methods for the minimization problem

The development of effective algorithms for the solution of optimization problems has been an active area of research for many years (see, e.g., [11]). The optimization problem that arises in our formulation at every time step can be posed either as a general minimization problem (3.3) or, recognizing some of its special structure, as a nonlinear least squares problem with zero residual (3.1). Since the saturation variables are physically constrained to lie in the closed interval [0,1], we need to handle the above formulations in the presence of simple bounds.

A myriad of algorithms for these problems have appeared in the literature. For the least squares problem, we apply the implementation found in the Port Library [3] due to Dennis and Gay (see [9], [10], and [13]). This implementation is well tested, robust and employs many modern heuristics to enhance and improve the numerical performance. The calculation of search directions [13] includes an adaptive scaling that keeps relative magnitudes of derivative information in a manageable range.

This implementation uses a trust-region globalization based on approximating the Levenberg-Marquardt parameter. The trust region also handles the simple bounds causing us to make no *ad hoc* extensions of the constitutive relationships such as the capillary pressure relationship to saturations that are negative or greater than one. Even in the neighborhood of the simple bounds, the globalization in the test problems discussed in the next section always became inactive when iterates drew close to the solution. This serendipitous behavior is due to the fact that the full gradient, not just the projected gradient, is zero at the solution since the residual is zero at the solution.

For the minimization formulation, we use the BFGS [6, 12, 15, 21] secant method also in the Port Library with the same trust-region globalization to handle the simple bounds. In both methods, no attempt was made to fine tune the input parameters.

5. Numerical experiments and conclusions

We consider a quarter five spot oil-water system. The capillary pressure curve is shown on Figure 1, and the relative permeabilities are $k_i(s_i) = s_i^2$. The initial water saturation is $s_w = 0.0$ and $s_w = 0.4$ in the two tests. For the other parameter values see Table 1. Both wells have a specified bottom hole pressure. A standard well model (see, e.g., [17], p. 79) with the Peaceman correction [18] is used. The oil pressure and velocity at the end of one time step are depicted on Figure 2 for the $s_w = 0$ problem.

Results are reported using both the identity operator and $A_n(S) + A_w(S)$ as the weighting operator W. The number of quasi-Newton steps and total number of trust-region evaluations needed to take one time step are reported in Table 2. In Table 2, "L-M" denotes the Levenberg-Marquardt method, and "BFGS" denotes the secant method with BFGS update found in the Port Library [3]. In both methods the trust-region strategy due to Gay [14] is used for globalization and to enforce the simple bounds $0 \leq S \leq 1$. The notation "DNC" is used if the method did not converge in less than 1000 trust-region evaluations. Since the least squares objective function has a zero residual at the solution, the results for two tolerances are reported, a liberal tolerance often used in practice of a 10^{-6} reduction in the relative residual and a more demanding 10^{-12} reduction to test robustness. A relative residual reduction of 10^{-6} corresponds to a mass balance error of approximately one percent, and a reduction of 10^{-12} corresponds to a mass balance error of approximately 10^{-10}.

As we can see from Table 2, the use of the weighting operator $W = A_w + A_n$ does improve convergence. However, even with the weighting function, the BFGS secant method is unable to solve the resulting minimization problem in a reasonable number of function evaluations. While the Levenberg-Marquardt method converges in a reasonable number of iterations, it uses the Jacobian matrix each iteration to calculate a step. Since calculation of the Jacobian matrix requires 100 gradient evaluations on our 10 by 10 mesh, this is not practical.

Since the residual is zero at the solution, the Levenberg-Marquardt method is essentially Gauss-Newton which in turn is essentially Newton's method applied to the following nonlinear system of equations:

$$A_n(S)P_n = q_n(S), \tag{5.1}$$

$$A_w(S)P_w = q_w(S), \tag{5.2}$$

$$P_c(S) = P_n - P_w. \tag{5.3}$$

In particular, since Newton's method preserves linear relationships and the conservation equations are linear with respect to P_w and P_n, Gauss-Newton on the least squares problem is almost Newton on (5.1)–(5.3). The marginal difference comes from the fact that the Gauss-Newton method neglects the part of the Hessian of the form

$$\sum_{i=1}^{n} R_i \cdot \nabla^2 R_i$$

where R_i is the i^{th} residual. Note that this part of the Hessian vanishes at the solution since R_i is zero at the solution. In light of this fact, it is hard to imagine that a Newton or inexact Newton method for (5.1)–(5.3) would not be computationally more competitive than Gauss-Newton on the least squares problem.

REFERENCES

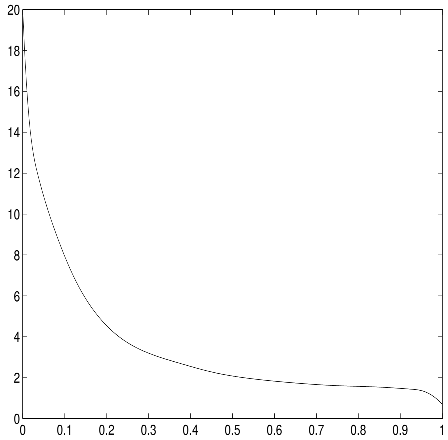

Fig. 1: Capillary pressure curve

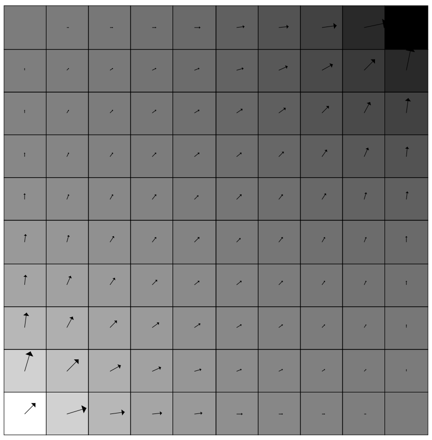

Fig. 2: Oil pressure and velocity

dt	0.1 days
LX	40. ft
LY	40. ft
dx	4. ft
dy	4. ft
Initial P_n	250. psia
Production Well Pressure	150. psia
Injection Well Pressure	500. psia
Production Well Radius	.4 ft
Injection Well Radius	.4 ft
Porosity	0.28
Permeability	1. md
Oil Viscosity	1.6 cp
Oil Density	48. lb/ft^3
Water Viscosity	0.23 cp
Water Density	62.34 lb/ft^3

TABLE 1: Parameters for Test Problem

[1] T. ARBOGAST, C. N. DAWSON, P. T. KEENAN, M. F. WHEELER, AND I. YOTOV, *Implementation of mixed finite element methods for elliptic equations on general geometry*. Submitted.

[2] T. ARBOGAST, M. F. WHEELER, AND I. YOTOV, *Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences*. To appear in SIAM J. Numer. Anal.

[3] AT&T BELL LABORATORIES, *Port 3: Mathematical Subroutine Library*, 1984.

[4] K. AZIZ AND A. SETTARI, *Petroleum reservoir simulation*, Applied Science Publishers, London and New York, 1979.

[5] F. BREZZI AND M. FORTIN, *Mixed and hybrid finite element methods*, Springer-Verlag, New York, 1991.

[6] C. G. BROYDEN, *The convergence of a class of double-rank minimization algorithms*, JIMA, 6 (1970), pp. 76–90, 222–236.

[7] G. CHAVENT AND J. JAFFRE, *Mathematical models and finite elements for reservoir simulation*, North-Holland, Amsterdam, 1986.

[8] Z. CHEN, *Expanded mixed finite element methods for linear second order elliptic problems i*, in IMA preprint series # 1219, Institute for Mathematics and its Applications, University of Minnesota, 1994.

[9] J. E. DENNIS, D. M. GAY, AND R. E. WELSCH, *An adaptive nonlinear least-squares algorithm*, ACM Transactions on Mathematical Software, 7 (1981), pp. 348–368.

[10] ———, *Algorithm 573: NL2SOL — an adaptive nonlinear least-squares algorithm [e4]*, ACM Transactions on Mathematical Software, 7 (1981), pp. 369–383.

[11] J. E. DENNIS AND R. B. SCHNABEL, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice–Hall, Englewood Cliffs, New Jersey, 1983.

[12] R. FLETCHER, *A new approach to variable metric algorithms*, Comput. J., 13 (1970), pp. 317–322.

[13] D. M. GAY, *Computing optimal locally constrained steps*, SIAM J. Sci. Stat. Comp., 2 (1981), pp. 186–197.

[14] ———, *Algorithm 611: Subroutines for unconstrained minimization using a model/trust-region approach*, ACM Transactions on Mathematical Software, 9 (1983), pp. 503–524.

[15] D. GOLDFARB, *A family of variable metric methods derived by variational means*, Math. Comp., 30 (1970), pp. 23–26.

[16] J. KOEBBE, *A computationally efficient modification of mixed finite element methods for flow problems with full transmissivity tensors*, Numer. Meth. for PDE's, 9 (1993), pp. 339–355.

[17] C. MATTAX AND R. DALTON, *Reservoir Simulation*, vol. 13, SPE–Monograph Series, Richardson, TX, 1990.

[18] D. W. PEACEMAN, *Interpretation of well-block pressures in numerical reservoir simulation*, SPEJ, June (1978), pp. 183–194.

[19] R. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics, vol. 606, Springer-Verlag,

Method	10^{-6} reduction		10^{-12} reduction	
	Quasi-Newton Steps	Total Steps	Quasi-Newton Steps	Total Steps
L-M, $S = 0$, $W = I$	4	7	9	12
L-M, $S = 0.4$, $W = I$	6	7	8	9
L-M, $S = 0$, $W = A_w + A_n$	4	7	7	10
L-M, $S = 0.4$, $W = A_w + A_n$	3	6	4	7
BFGS, $S = 0$, $W = I$	DNC	DNC	DNC	DNC
BFGS, $S = 0.4$, $W = I$	253	554	DNC	DNC
BFGS, $S = 0$, $W = A_w + A_n$	173	265	DNC	DNC
BFGS, $S = 0.4$, $W = A_w + A_n$	44	61	439	627

TABLE 2: Results for Test Problem

New York, 1977, pp. 292–315.

[20] T. F. RUSSELL AND M. F. WHEELER, *Finite element and finite difference methods for continuous flows in porous media*, in The Mathematics of Reservoir Simulation, R. E. Ewing, ed., SIAM, Philadelphia, 1983, pp. 35–106.

[21] D. F. SHANNO, *Conditioning of quasi-newton methods for function minimization*, Math. Comp., 24 (1970), pp. 647–657.

[22] M. F. WHEELER, K. R. ROBERSON, AND A. CHILAKAPATI, *Three-dimensional bioremediation modeling in heterogeneous porous media*, in Computational Methods in Water Resources IX, Vol. 2: Mathematical Modeling in Water Resources, T. F. Russell, R. E. Ewing, C. A. Brebbia, W. G. Gray, and G. F. Pindar, eds., Southampton, U. K., 1992, Computational Mechanics Publications, pp. 299–315.