JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 111, No. 3, pp. 473–488, December 2001 (©2001)

New Optimization Approach to Multiphase Flow[1]

A. J. Kearsley,[2] L. C. Cowsar,[3] R. Glowinski,[4]
M. F. Wheeler,[5] and I. Yotov[6]

Abstract. A new optimization formulation for simulating multiphase flow in porous media is introduced. A locally mass-conservative, mixed finite-element method is employed for the spatial discretization. An unconditionally stable, fully-implicit time discretization is 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 quasi-Newton 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.

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

[1]This work was supported in part by the Department of Energy, the National Science Foundation, and the State of Texas Governor's Energy Office.

[2]Assistant Professor, Department of Mathematical Science, Carnegie Mellon University, Pittsburgh, Pennsylvania.

[3]Staff Member, Lucent Technologies, Murray Hill, New Jersey.

[4]Professor, Department of Mathematics, University of Houston, Houston, Texas.

[5]Professor, Texas Institute of Computational and Applied Mathematics, University of Texas at Austin, Austin, Texas.

[6]Assistant Professor, Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania.

0022-3239/01/1200-0473$19.50/0

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 requires 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 requires only 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.

In this work, we consider an immiscible two-phase flow in a porous medium. Let $\Omega \subset R^d$, $d = 2, 3$, be the domain of the problem. The equations governing the fluid motion through a porous medium are the mass conservation of the phases (Refs. 1–2),

$$\partial(\phi s_i \rho_i)/\partial t + \nabla \cdot \rho_i u_i = q_i, \qquad (x, t) \in \Omega \times (0, T), \tag{1}$$

where $i = w$ (wetting) or $i = n$ (nonwetting) denotes the phase, s_i is the phase saturation, ρ_i is the phase density, ϕ is the porosity, q_i is the source term, and

$$u_i = -[k_i(s_i)K/\mu_i](\nabla p_i - \rho_i g \nabla D), \qquad (x, t) \in \Omega \times (0, T), \tag{2}$$

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,

$$s_w + s_n = 1, \tag{3}$$

and the capillary pressure relation,

$$p_c(s_w) = p_n - p_w. \tag{4}$$

The no-flow boundary conditions

$$u_i \cdot \nu = 0, \qquad (x, t) \in \partial\Omega \times (0, T), \tag{5}$$

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

$$s_w(x, 0) = s_0(x), \qquad x \in \Omega, \tag{6}$$

complete the model. We note that the no-flow boundary conditions are assumed only for the 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 assume also 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

$$k_i(s_i) = 0, \qquad \text{for } s_i = 0,\ i = w, n.$$

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 determines uniquely 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)–(2) with (3) and (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

$$\begin{aligned} f(s) &= (1/2)\|F(s)\|^2_{L^2(\Omega)} \\ &\equiv (1/2)\int_\Omega |F(s(x))|^2\,dx, \end{aligned} \tag{7}$$

and consider the unconstrained 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)–(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 (Refs. 3–7) 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-difference and mixed methods were first noticed by Russell and Wheeler (Ref. 8) and extends to the expanded mixed method (Ref. 6).

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

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

with the norm

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

and set

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

Following Ref. 6, we introduce the pressure gradients

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

Then,

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

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 $u_i(\cdot, t) \in V$, $\tilde{u}_i(\cdot, t) \in (L^2(\Omega))^d$, $p_i(\cdot, t) \in L^2(\Omega)$, $i = w, n$, and $s(\cdot, t) \in L^2(\Omega)$ such that

$$(u_i, \tilde{v}) = (\lambda_i(s) K \tilde{u}_i, \tilde{v}), \qquad \tilde{v} \in (L^2(\Omega))^d, \tag{8}$$

$$(\tilde{u}_i, v) - (p_i, \nabla \cdot v) = 0, \qquad v \in V, \tag{9}$$

$$\phi(\partial s/\partial t, w) + (\nabla \cdot u_w, w) = (\tilde{q}_w(s), w), \qquad w \in L^2(\Omega), \tag{10}$$

$$-\phi(\partial s/\partial t, w) + (\nabla \cdot u_n, w) = (\tilde{q}_n(s), w), \qquad w \in L^2(\Omega), \tag{11}$$

$$(p_n - p_w, w) = (p_c(s), w), \qquad w \in L^2(\Omega), \tag{12}$$

$$(s(\cdot, 0), w) = (s_0, w), \qquad w \in L^2(\Omega), \tag{13}$$

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 $V_h \times W_h \subset V \times L^2(\Omega)$ be any of the known mixed finite-element spaces on $\mathcal{T}_h$; see Ref. 9 for a detailed description of these spaces. Let $\tilde{V}_h$ *be a finite element subspace of* $(L^2(\Omega))^d$ *such that* $V_h \subseteq \tilde{V}_h$. 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 $0 \le n \le N$, $(u_{i,h}^n, \tilde{u}_{i,h}^n, p_{i,h}^n) \in V_h \times \tilde{V}_h \times W_h$ and $s_h^n \in W_h$ such that

$$(u_{i,h}^n, \tilde{v}) = (\lambda_i(s_h^n) K \tilde{u}_{i,h}^n, \tilde{v}), \qquad \tilde{v} \in \tilde{V}_h, \tag{14}$$

$$(\tilde{u}_{i,h}^n, v) - (p_{i,h}^n, \nabla \cdot v) = 0, \qquad v \in V_h, \tag{15}$$

$$\phi((s_h^n - s_h^{n-1})/\Delta t^n, w) + (\nabla \cdot u_{w,h}^n, w) = (\tilde{q}_w(s_h^n), w), \qquad w \in W_h, \tag{16}$$

$$-\phi((s_h^n - s_h^{n-1})/\Delta t^n, w) + (\nabla \cdot u_{n,h}^n, w) = (\tilde{q}_n(s_h^n), w), \qquad w \in W_h, \tag{17}$$

$$(p_{n,h}^n - p_{w,h}^n, w) = (p_c(s_h^n), w), \qquad w \in W_h, \tag{18}$$

$$(s_h^0, w) = (s_0, w), \qquad w \in W_h. \tag{19}$$

2.2. Cell-Centered Finite-Difference Scheme. To obtain a finite-difference method, we consider the lowest-order Raviart–Thomas spaces (Ref. 10) and take $\tilde{V}_h = V_h$ 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 Ref. 11. Let

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

be the velocity space, and let

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

be the pressure space. Let

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

We employ quadrature rules to approximate the vector integrals in (14)–(15). The two equations are replaced by

$$(u^n_{i,h}, \tilde{v})_T = (\lambda_i(s^n_h) K \tilde{u}^n_{i,h}, \tilde{v})_T, \qquad \tilde{v} \in \tilde{V}_h, \tag{20}$$

$$(\tilde{u}^n_{i,h}, v)_T - (p^n_{i,h}, \nabla \cdot v) = 0, \qquad v \in V_h, \tag{21}$$

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 (20)–(21) for the rest of this section. The application of the trapezoidal rule leads to diagonal coefficient matrices for u_h and $\tilde{u}_h$ in (20)–(21). Therefore, u_h and $\tilde{u}_h$ 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

$$(x_{i+1/2}, y_{j+1/2}), \qquad i = 0, \ldots, N_x \text{ and } j = 0, \ldots, N_y,$$

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

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

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

Let us write $v = (v^x, v^y)$ for $v \in R^2$; 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.

For simplicity, let us assume that K is diagonal; see Ref. 6 for the full tensor case. Let

$$\mathcal{K}(s) = \lambda(s) K.$$

Choosing $\tilde{v}$ in (20) to be the basis function at node $(i+1/2, j)$ or at node $(i, j+1/2)$, we get

$$u^x_{h,i+1/2,j} = (\mathscr{K}_{11})_{i+1/2,j}\tilde{u}^x_{h,i+1/2,j}, \tag{22a}$$

$$u^y_{h,i,j+1/2} = (\mathscr{K}_{22})_{i,j+1/2}\tilde{u}^y_{h,i,j+1/2}, \tag{22b}$$

which is a finite-difference approximation of $u = \mathscr{K}\tilde{u}$. The same choice of v in (21) gives

$$\tilde{u}^x_{h,i+1/2,j} = -(p_{h,i+1,j} - p_{h,i,j})/(1/2)(h^x_i + h^x_{i+1}), \tag{23a}$$

$$\tilde{u}^y_{h,i,j+1/2} = -(p_{h,i,j+1} - p_{h,i,j})/(1/2)(h^y_j + h^y_{j+1}), \tag{23b}$$

which is a finite-difference approximation of $\tilde{u} = -\nabla p$. Finally, choosing w in (16)–(17) to be the basis function at element (i, j), we get

$$\{\phi((s^n_{h,i,j} - s^{n-1}_{h,i,j})/(\Delta t^n)) + (u^{x,n}_{w,h,i+1/2,j} - u^{x,n}_{w,h,i-1/2,j})/h^x_i + (u^{y,n}_{w,h,i,j+1/2} - u^{y,n}_{w,h,i,j-1/2})/h^y_j\}h^x_i h^y_j \int_{E_{ij}} \tilde{q}_w(s^n_h)\,dx\,dy, \tag{24}$$

with a similar expression for (17).

The combination of (22)–(24) produces the cell-centered finite-difference approximation of

$$\phi(\partial s_i/\partial t) - \nabla\cdot\mathscr{K}_i(s_i)\nabla p_i = \tilde{q}_i, \qquad i = w, n.$$

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

$$(\mathscr{K}_{11})_{i+1/2,j} = \lambda(s_h)_{i+1/2,j}(K_{11})_{i+1/2,j},$$

Equation (22) 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 computed directly. One-point upstream weighting is used to determine this value, i.e.,

$$\lambda(s_h)_{i+1/2,j} = \lambda(s_h)_{i,j} \quad \text{or} \quad \lambda(s_h)_{i+1/2,j} = \lambda(s_h)_{i+1,j},$$

depending on the direction of the flow. If K is discontinuous and is given as a piecewise constant function, the harmonic average is used for its value on the edges, i.e.,

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

Remark 2.2. The cell-centered finite-difference scheme described above has been used commonly by the environmental and petroleum engineers (Ref. 12). 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. 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

$$F(s_h^n) = p_c(s_h^n) - (p_n(s_h^n) - p_w(s_h^n)), \tag{25}$$

where $p_n(s_h^n)$ and $p_w(s_h^n)$ are the solution to the discrete system (20)–(21), (16)–(17) 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: $W_h \to R$ by

$$f(s_h) = (1/2)\int_\Omega (W(s_h)F(s_h))F(s_h)\,dx. \tag{26}$$

It is easy to see that solving the discrete system (16)–(18), (20)–(21) for s_h^n is equivalent to finding s_h such that $f(s_h) = 0$ or to solving the minimization problem

$$\min_{s_h \in W_h} f(s_h). \tag{27}$$

Remark 3.1. If the identity operator is used for the weighting operator $W(s)$, then (26) is the discrete analog of (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 (20)–(21), (16)–(17) in the form

$$A_i(S)P_i(S) = q_i(S), \tag{28}$$

where S is the vector of the saturation unknowns, P_i is the vector of pressure 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 (28) for $i = n, w$ and using the fact that

$$P_c(S) = P_n(S) - P_w(S)$$

at the solution due to (18), we see that

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

Since the saturation of the wetting phase and the saturation of the nonwetting 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 semidefinite in general. Adding Eq. (28) for $i = n$, we arrive at a nonsingular system that defines $P_n(S)$, namely,

$$(2A_n(S) + A_w(S))P_n(S) = 2q_n(S) + q_w(S) + A_w(S)P_c(S). \tag{29}$$

A similar relationship holds for P_w; namely,

$$(A_n(S) + 2A_w(S))P_w(S) = q_n(S) + 2q_w(S) - A_n(S)P_c(S). \tag{30}$$

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 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 (27) are described in the next section. They require the 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 (29)–(30). 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,

$$f(S) = f(S, P_n(S), P_w(S)). \tag{31}$$

By the chain rule, we have

$$df/dS = \partial f/\partial S + (\partial f/\partial P_n)(dP_n/dS) + (\partial f/\partial P_w)(dP_w/dS). \tag{32}$$

Differentiating (28), we find that dP_i/dS satisfies

$$A_i(dP_i/dS) = dq_i/dS - (dA_i/dS)P_i, \tag{33}$$

an equation involving the same operator as the evaluation of the objective function. Combining (32)–(33) we get

$$\begin{aligned} df/dS = {} & \partial f/\partial S + (\partial f/\partial P_n)A_n^{-1}[dq_n/dS - (dA_n/dS)P_n] \\ & + (\partial f/\partial P_w)A_w^{-1}[dq_w/dS - (dA_w/dS)P_w]. \end{aligned} \tag{34}$$

Therefore,

$$\begin{aligned} \nabla f = {} & (\partial f/\partial S)^T + [dq_n/dS - (dA_n/dS)P_n]^T A_n^{-T}(\partial f/\partial P_n)^T \\ & + [dq_w/dS - (dA_w/dS)P_w]^T A_w^{-T}(\partial f/\partial P_w)^T. \end{aligned} \tag{35}$$

Since the stiffness matrices A_i are symmetric,

$$A_i^{-T} = A_i^{-1}.$$

The partial derivative in (35) can be computed easily by the formulas

$$\partial f/\partial S = (1/2)F^T(dW/dS)F + (WF)^T(dP_c/dS),$$

$$\partial f/\partial P_n = -(WF)^T,$$

$$\partial f/\partial P_w = (WF)^T.$$

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 unconstrained optimization problems has been an active area of research for many years; see e.g. Ref. 14. The optimization problem that arises in our formulation at every time step can be posed either as a general minimization

problem (27) or, recognizing some of its special structure, as a nonlinear least squares problem with zero residual (25). 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 (Ref. 15) due to Dennis and Gay (see Refs. 16–18). This implementation is well tested, robust, and employs many modern heuristics to enhance and improve the numerical performance. The calculation of search directions (Ref. 18) includes an adaptive scaling that keeps the relative magnitudes of the derivative information in a manageable range.

This implementation uses a trust-region globalization based on approximating the Levenberg–Marquardt parameter. The trust region handles also 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 became always 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 (Refs. 19–22) 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 Fig. 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., Ref. 12, p. 79) with the Peaceman correction (Ref. 13) is used. The oil pressure and velocity at the end of one time step are depicted on Fig. 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, LM denotes the Levenberg–Marquardt method, and BFGS denotes the secant method with BFGS update

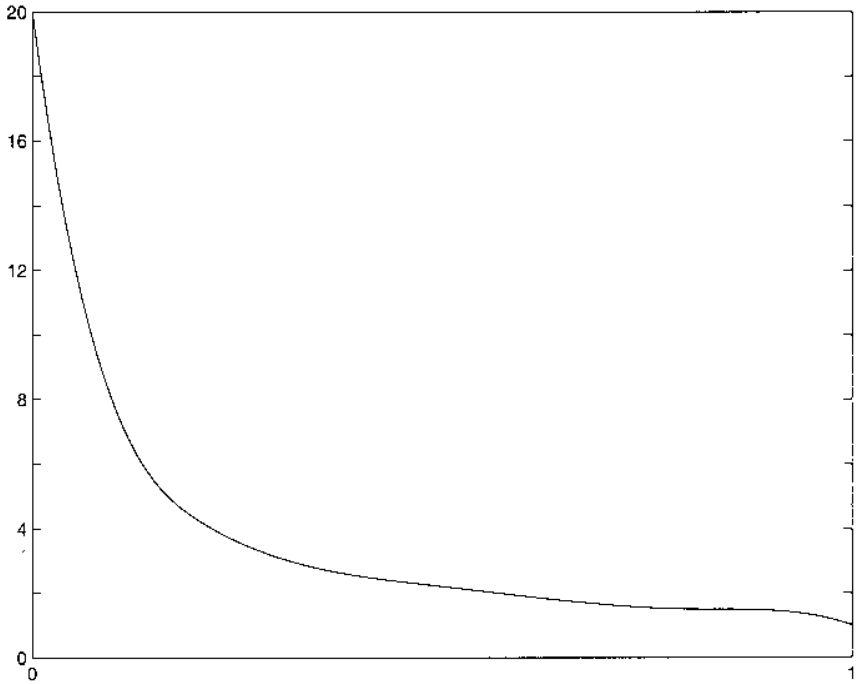

Fig. 1. Capillary pressure curve.

Table 1. Parameters for test problem.

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	0.4 ft
Injection well radius	0.4 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

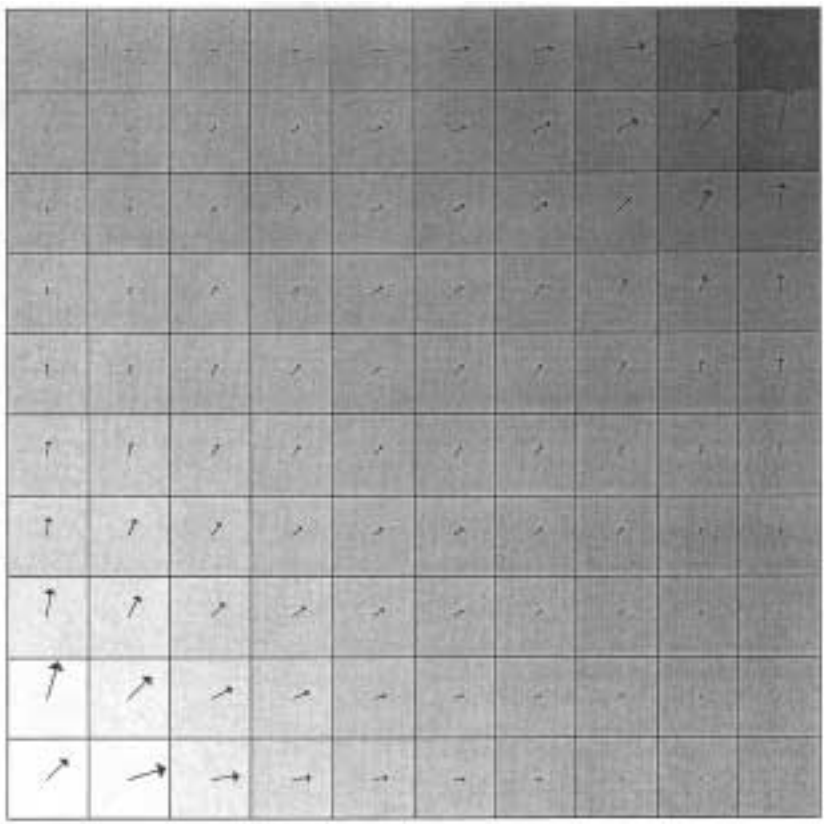

Fig. 2. Oil pressure and velocity.

Table 2. Results for test problem.

Method			10^{-6} Reduction		10^{-12} Reduction	
			Quasi-Newton steps	Total steps	Quasi-Newton steps	Total steps
LM,	$S = 0$,	$W = I$	4	7	9	12
LM,	$S = 0.4$,	$W = I$	6	7	8	9
LM,	$S = 0$,	$W = A_w + A_n$	4	7	7	10
LM,	$S = 0.4$,	$W = A_w + A_n$	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

found in the Port Library (Ref. 15). In both methods, the trust-region strategy due to Gay (Ref. 23) 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 at 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 the Gauss–Newton, which in turn is essentially the Newton method applied to the following nonlinear system of equations:

$$A_n(S)P_n = q_n(S), \tag{36}$$

$$A_w(S)P_w = q_w(S), \tag{37}$$

$$P_c(S) = P_n - P_w. \tag{38}$$

In particular, since the Newton method preserves linear relationships, and since the conservation equations are linear with respect to P_w and P_n, the Gauss–Newton method on the least squares problem is almost the Newton method on (36)–(38). 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 ith 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 (36)–(38) would not be computationally more competitive than the Gauss–Newton on the least squares problem.

References

1. Aziz, K., and Settari, A., *Petroleum Reservoir Simulation*, Applied Science Publishers, London, England, 1979.
2. Chavent, G., and Jaffre, J., *Mathematical Models and Finite Elements for Reservoir Simulation*, North-Holland, Amsterdam, Holland, 1986.
3. Wheeler, M. F., Roberson, K. R., and Chilakapati, A., *Three-Dimensional Bioremediation Modeling in Heterogeneous Porous Media*, Mathematical Modeling in Water Resources, Edited by T. F. Russel, R. E. Ewing, C. A. Brebbia, W. G. Gray, and G. F. Pinder, Computational Mechanics Publications, Southampton, England, Vol. 2, pp. 299–315, 1992.
4. Koebbe, J., *A Computationally Efficient Modification of Mixed Finite-Element Methods for Flow Problems with Full Transmissivity Tensors*, Numerical Methods for Partial Differential Equations, Vol. 9, pp. 339–355, 1993.
5. Chen, Z., *Expanded Mixed Finite-Element Methods for Linear Second-Order Elliptic Problems, I*, Technical Report 1219, Institute for Mathematics and Its Applications, University of Minnesota, 1994.
6. Arbogast, T., Wheeler, M. F., and Yotov, I., *Mixed Finite Elements for Elliptic Problems with Tensor Coefficients as Cell-Centered Finite Differences*, SIAM Journal on Numerical Analysis, Vol. 34, pp. 828–852, 1997.
7. Yotov, I., *Mixed Finite-Element Methods for Flow in Porous Media*, PhD Thesis, Rice University, 1996.
8. Russel, T. F., and Wheeler, M. F., *Finite-Element and Finite-Difference Methods for Continuous Flows in Porous Media*, Mathematics of Reservoir Simulation, Edited by R. E. Ewing, SIAM, Philadelphia, Pennsylvania, pp. 35–106, 1983.
9. Brezzi, F., and Fortin, M., *Mixed and Hybrid Finite-Element Methods*, Springer Verlag, New York, NY, 1991.
10. Raviart, P. A., and Thomas, J. M., *A Mixed Finite-Element Method for Second-Order Elliptic Problems*, Finite-Element Methods, Edited by I. Galligani and E. Magenes, Lecture Notes in Mathematics, Springer Verlag, Berlin, Germany, Vol. 606, pp. 292–315, 1977.
11. Arbogast, T., Dawson, C. N., Keenan, P. T., Wheeler, M. F., and Yotov, I., *Enhanced Cell-Centered Finite Differences for Elliptic Equations on General Geometry*, SIAM Journal on Scientific and Statistical Computing, Vol. 19, pp. 404–425, 1997.
12. Mattax, C., and Dalton, R., *Reservior Simulation*, Society of Petroleum Engineers, Monograph Series, Vol. 13, 1990.
13. Peaceman, D. W., *Interpretation of Well-Block Pressures in Numerical Reservoir Simulation*, Society of Petroleum Engineers Journal, pp. 183–194, 1978.
14. Dennis, J. E., Jr., and Schnabel, R. B., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, New Jersey, 1983.
15. ANONYMOUS, *Part 3 Library: Mathematical Subroutine Library*, ATT Bell Laboratories, Murray Hill, New Jersey, 1984.

16. DENNIS, J. E., GAY, D. M., and WELSCH, R. E., *An Adaptive Nonlinear Least-Squares Algorithm*, ACM Transactions on Mathematical Software, Vol. 7, pp. 348–368, 1981.
17. DENNIS, J. E., GAY, D. M., and WELSCH, R. E., *Algorithm 573: NL2SOL, An Adaptive Nonlinear Least-Squares Algorithm*, ACM Transactions on Mathematical Software, Vol. 7, pp. 369–383, 1981.
18. GAY, D. M., *Computing Optimal Locally Constrained Steps*, SIAM Journal on Scientific and Statistical Computing, Vol. 2, pp. 186–197, 1981.
19. BROYDEN, C. G., *The Convergence of a Class of Double-Rank Minimization Algorithms*, Journal of the Institute of Mathematics and Its Applications, Vol. 6, pp. 76–90, 1970.
20. FLETCHER, R., *A New Approach to Variable-Metric Algorithms*, Computer Journal, Vol. 13, pp. 317–322, 1970.
21. GOLDFARB, D., *A Family of Variable-Metric Methods Derived by Variational Means*, Mathematics of Computation, Vol. 30, pp. 23–26, 1970.
22. SHANNO, D. F., *Conditioning of Quasi-Newton Methods for Function Minimization*, Mathematics of Computation, Vol. 24, pp. 647–657, 1970.
23. GAY, D. M., *Algorithm 611: Subroutines for Unconstrained Minimization Using a Model Trust-Region Approach*, ACM Transactions on Mathematical Software, Vol. 9, pp. 503–524, 1983.