WELL-CONDITIONED ITERATIVE SCHEMES FOR MIXED FINITE-ELEMENT MODELS OF POROUS-MEDIA FLOWS

Myron B. Allen
Richard E. Ewing
Peng Lu

1992
Journal Article
WWRC-92-12

In

SIAM: Journal of Scientific
and Statistical Computing

Volume 13, Number 3

Submitted by

Myron B. Allen
Richard E. Ewing
Department of Mathematics
University of Wyoming
Laramie, Wyoming

Peng Lu
Department of Mathematics
University of Georgia
Athens, Georgia

SIAM J. SCI. STAT. COMPUT.
Vol. 13, No. 3, pp. 794–814, May 1992

WELL-CONDITIONED ITERATIVE SCHEMES FOR MIXED FINITE-ELEMENT MODELS OF POROUS-MEDIA FLOWS*

MYRON B. ALLEN†, RICHARD E. EWING†, AND PENG LU‡

Abstract. Mixed finite-element methods are attractive for modeling flows in porous media since they can yield pressures and velocities having comparable accuracy. In solving the resulting discrete equations, however, poor matrix conditioning can arise both from spatial heterogeneity in the medium and from the fine grids needed to resolve that heterogeneity. This paper presents two iterative schemes that overcome these sources of poor conditioning. The first scheme overcomes poor conditioning resulting from the use of fine grids. The idea behind the scheme is to use spectral information about the matrix associated with the discrete version of Darcy's law to precondition the velocity equations, employing a multigrid method to solve mass-balance equations for pressure or head. This scheme still exhibits slow convergence when the permeability or hydraulic conductivity is highly variable in space. The second scheme, based on the first, uses mass lumping to precondition the Darcy equations, thus requiring more work per iteration and minor modifications to the multigrid algorithm. However, the scheme is insensitive to heterogeneities. The overall approach should also be useful in such applications as electric field simulation and heat transfer modeling when the media in question have spatially variable material properties.

Key words. mixed finite elements, iterative solution schemes, heterogeneous porous media

AMS(MOS) subject classification. 65

1. Introduction.

We consider methods for solving discrete approximations to the equations governing single-fluid flow in a porous medium. If the flow is steady and two-dimensional with no gravity drive, Darcy's law and the mass balance take the following forms:

$$\begin{aligned} \mathbf{u} &= -K \operatorname{grad} p \quad \text{in } \Omega, \\ \operatorname{div} \mathbf{u} &= f \quad \text{in } \Omega. \end{aligned} \tag{1.1}$$

Here $\mathbf{u}, p$, and f represent the Darcy velocity, pressure, and source term, respectively. For simplicity, we take the spatial domain to be a square, scaled so that $\Omega = (0,1) \times (0,1)$. The coefficient $K(x,y)$ is the *mobility*, defined as the ratio of the permeability of the porous medium to the dynamic viscosity of the fluid. In applications to underground flows, the structure of K may be quite complex, depending on the lithology of the porous medium and the composition of the fluid. We assume, however, that this ratio is bounded and integrable on $\bar{\Omega}$ and satisfies $K \geq K_{\text{inf}} > 0$. We impose the boundary condition $p = 0$ on $\partial\Omega$, so that p effectively represents the deviation in pressure from a reference value known along $\partial\Omega$.

Scientists modeling contaminant flows in groundwater or solvent flows in oil reservoirs often need accurate finite-element approximations of $\mathbf{u}$ and p simultaneously. For this reason, mixed finite-element methods for solving the system (1.1) are particularly attractive, since they can yield approximations to $\mathbf{u}$ and p that have comparable accuracy [1], [5], [9]. The key to achieving such approximations is the use of appropriate piecewise polynomial trial spaces, such as those proposed by

* Received by the editors April 12, 1990; accepted for publication (in revised form) February 20, 1991. This research was supported by National Science Foundation grant RII-8610680, Office of Naval Research grant 0014-88-K-0370, and the Wyoming Water Reseach Center.

† Department of Mathematics, University of Wyoming, Laramie, Wyoming 82071-3036.

‡ Department of Mathematics, University of Georgia, Athens, Georgia 30605.

Raviart and Thomas [11]. As we review in §2, if we use the lowest-degree Raviart–Thomas spaces, the mixed formulation yields systems of discrete equations that have the form

$$\begin{aligned} AU + NP &= 0, \\ N^T U &= F. \end{aligned} \tag{1.2}$$

Here, U and P signify vectors containing nodal values of the trial functions for $\mathbf{u}$ and p, defined on a grid over Ω, and A and N are matrices. As we illustrate below, the matrix A contains information about the spatially varying material property K, while N and N^T are essentially finite-difference matrices.

Equations (1.2) can be quite difficult to solve efficiently, for the following reasons. When K varies over short distances, accurate finite-element approximations require fine grids on Ω. For example, one might choose grids fine enough to allow reasonable approximations of K by piecewise constant functions. Fine grids, however, typically yield poorly conditioned matrix equations. For classical stationary iterative schemes, this increase in the condition number of the system leads to slow convergence, no matter how "nice" K may be [2, §4.11]. The problem is compounded whenever K exhibits large spatial variations, as can occur near lithologic changes in the porous medium or sharp contacts between fluids of different viscosity. In such problems, as we shall demonstrate, the poor conditioning associated with spatial variability typically aggravates that associated with the fine grids needed to resolve the physics of the problem. Thus, in problems with significant material heterogeneity, methods that are relatively insensitive to these two sources of poor conditioning can have considerable utility.

In this paper we discuss two iterative schemes for the mixed-method equations (1.2). The first scheme possesses convergence rates that are independent of the fineness of the grid. The second scheme, derived from the first, also overcomes the sensitivity to the spatial structure of K, at the expense of somewhat more computation per iteration. Briefly, the first scheme proceeds as follows: Let $(U^{(0)}, P^{(0)})$ be initial guesses for the value of (U, P). Then the kth iterate for (U, P) is the solution of

$$\begin{pmatrix} \omega I & N \\ N^T & 0 \end{pmatrix} \begin{pmatrix} U^{(k)} \\ P^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix} + \begin{pmatrix} \omega I - A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U^{(k-1)} \\ P^{(k-1)} \end{pmatrix}, \tag{1.3}$$

where I stands for the identity matrix and ω signifies a parameter, discussed below, that is related to the spectral radius $\rho(A)$ of A. For each iteration level k, the main computational work in (1.3) is to solve a linear system of the form $(\omega^{-1} N^T N) P^{(k)} = G^{(k-1)}$. However, the matrix $\omega^{-1} N^T N$ remains vulnerable to the poor conditioning associated with fine grids. We overcome this difficulty by using a multigrid scheme to solve for $P^{(k)}$, thereby greatly reducing the computational work in each iteration.

An interesting feature of this approach is that $N^T N$ is essentially the matrix associated with the five-point difference approximation to the Laplace operator with Dirichlet boundary conditions. Hence, the multigrid portion of the scheme does not encounter the variable coefficient, and the algorithm is particularly simple. The price paid for this simplicity, as we shall see, is sensitivity to the poor conditioning associated with spatial variability.

To overcome this second source of trouble, we modify the first scheme to get new ones of the form

$$\begin{pmatrix} D & N \\ N^T & 0 \end{pmatrix} \begin{pmatrix} U^{(k)} \\ P^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix} + \begin{pmatrix} D - A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U^{(k-1)} \\ P^{(k-1)} \end{pmatrix}, \tag{1.4}$$

where D denotes a diagonal matrix that we compute from A. This new class of schemes requires us to invert $N^T DN$, which we again do using a multigrid method to preserve h-independence of the convergence rate. While the multigrid method must now accommodate spatially varying coefficients, the overall scheme possesses the advantage that its convergence rate is independent of the spatial structure of K, provided K is piecewise constant on the grids of interest.

Our paper has the following format. In §2 we review the mixed finite-element method that we use. Section 3 describes the first iterative scheme in more detail and analyzes its convergence. In §4 we discuss the application of multigrid ideas to the first scheme. Much of the motivation and groundwork for the second class of iterative schemes resides in §§3 and 4. In §5 we present some numerical results for this algorithm. Section 6 describes the modifications necessary to produce the second class of iterative schemes and presents numerical results illustrating good convergence rates even in the presence of heterogeneities.

2. A mixed finite-element method. We begin with a brief review of the mixed finite-element method, following the notation of Ewing and Wheeler [8]. Let $H(\text{div}, \Omega) = \{\mathbf{v} \in L^2(\Omega) \times L^2(\Omega) : \text{div } \mathbf{v} \in L^2(\Omega)\}$. The variational form for (1.1) is as follows: Find a pair $(\mathbf{u}, p) \in H(\text{div}, \Omega) \times L^2(\Omega)$ such that

$$\begin{aligned} &\int_\Omega \frac{\mathbf{u} \cdot \mathbf{v}}{K} dx\, dy - \int_\Omega p \text{ div } \mathbf{v}\, dx\, dy = 0 \quad \forall\, \mathbf{v} \in H(\text{div}, \Omega), \\ &\int_\Omega (\text{div } \mathbf{u} - f) q\, dx\, dy = 0 \quad \forall\, q \in L^2(\Omega). \end{aligned} \tag{2.1}$$

By our assumptions on K, there exist constants $K_{\text{inf}}, K_{\text{sup}}$ such that $0 < K_{\text{inf}} \leq K \leq K_{\text{sup}}$. Implicit in these equations is also the assumption that K^{-1} is integrable on $\bar{\Omega}$.

To discretize the system (2.1), let $\Delta_x = \{0 = x_0 < x_1 < \cdots < x_m = 1\}$ be a set of points on the x-axis and $\Delta_y = \{0 = y_0 < y_1 < \cdots < y_n = 1\}$ a set of points on the y-axis. Let $\Delta_h = \Delta_x \times \Delta_y$ be the rectangular grid on Ω with nodes $\{(x_i, y_j)\}_{i=0,j=0}^{m,n}$. The mesh of this grid is

$$h = \max_{i,j} \{x_i - x_{i-1},\, y_j - y_{j-1}\}.$$

We assume throughout the paper that Δ_x and Δ_y are quasi-uniform in the sense that $x_i - x_{i-1} \geq \alpha h$ and $y_j - y_{j-1} \geq \alpha h$ for some fixed $\alpha \in (0, 1)$. With Δ_h we associate a finite-element subspace $\mathbf{Q}_h \times V_h$ of $H(\text{div}, \Omega) \times L^2(\Omega)$. The "velocity space" is $\mathbf{Q}_h = Q_h^x \times Q_h^y$, where Q_h^x and Q_h^y are both tensor-product spaces of one-dimensional, finite-element spaces. In particular, we use the lowest-order Raviart–Thomas spaces in which Q_h^x contains functions that are piecewise linear and continuous on Δ_x and piecewise constant on Δ_y. Similarly, Q_h^y contains functions that are piecewise linear and continuous on Δ_y and piecewise constant on Δ_x. The "pressure space" V_h consists of functions that are piecewise constant on Δ_h.

Given these approximating spaces, the corresponding mixed finite-element method for solving (2.1) is as follows: Find a pair $(\mathbf{u}_h, p_h) \in \mathbf{Q}_h \times V_h$ such that

$$\begin{aligned} &\int_\Omega \frac{\mathbf{u}_h \cdot \mathbf{v}_h}{K} dx\, dy - \int_\Omega p_h \text{ div } \mathbf{v}_h dx\, dy = 0 \qquad \forall\, \mathbf{v}_h \in \mathbf{Q}_h, \\ &\int_\Omega (\text{div } \mathbf{u}_h - f) q_h dx\, dy = 0 \qquad\qquad\qquad \forall\, q_h \in V_h. \end{aligned} \tag{2.2}$$

This finite-element discretization yields approximations $\mathbf{u}_h$ and p_h whose global errors are both $O(h)$ in the norm $\|\cdot\|_{L^2(\Omega)}$. Ewing, Lazarov, and Wang [6] also prove superconvergence results that guarantee smaller errors at special points in Ω. This phenomenon appears in our numerical examples in §5. In contrast, standard approaches solve for approximations to p and then numerically differentiate to compute $\mathbf{u} = -K \operatorname{grad} p$, thereby losing an order of accuracy in the velocity field [1].

To see the linear algebraic equations implied by (2.2), suppose $\mathbf{u}_h$ and p_h have the expansions

$$\mathbf{u}_h(x,y) = \left(\sum_{i=0}^{m} \sum_{j=1}^{n} U_{i,j}^x \phi_{i,j}^x(x,y), \sum_{i=1}^{m} \sum_{j=0}^{n} U_{i,j}^y \phi_{i,j}^y(x,y) \right),$$

$$p_h(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} P_{i,j} \psi_{i,j}(x,y).$$

Here, $\phi_{i,j}^x$, $\phi_{i,j}^y$, and $\psi_{i,j}$ signify elements in the standard nodal bases for Q_h^x, Q_h^y, and V_h. Define the column vectors $U \in \mathbb{R}^{2mn+m+n}$, $P \in \mathbb{R}^{mn}$ containing the nodal unknowns as follows:

$$\begin{aligned} U^T = (&U_{0,1}^x, U_{1,1}^x, \cdots, U_{m,1}^x, \cdots, U_{0,n}^x, U_{1,n}^x, \cdots, U_{m,n}^x, \\ &U_{1,0}^y, U_{1,1}^y, \cdots, U_{1,n}^y, \cdots, U_{m,0}^y, U_{m,1}^y, \cdots, U_{m,n}^y), \end{aligned} \tag{2.3}$$

$$P^T = (P_{1,1}, P_{2,1}, \cdots, P_{m,1}, \cdots, P_{1,n}, P_{2,n}, \cdots, P_{m,n}).$$

Figure 1 shows how to associate these coefficients with nodes on a spatial grid Δ_h with $m = 4$, $n = 3$.

With these bases, the problem (2.2) has a matrix representation of the form

$$\begin{pmatrix} A & N \\ N^T & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix}. \tag{2.4}$$

Here A is a symmetric, positive definite matrix having the block structure

$$A = \begin{pmatrix} A^x & 0 \\ 0 & A^y \end{pmatrix},$$

in which $A^x \in \mathbb{R}^{(m+1)n \times (m+1)n}$ and $A^y \in \mathbb{R}^{m(n+1) \times m(n+1)}$ have entries of the form

$$\int_\Omega \frac{\phi_{i,j}^x \phi_{k,\ell}^x}{K} dx\, dy, \qquad \int_\Omega \frac{\phi_{i,j}^y \phi_{k,\ell}^y}{K} dx\, dy,$$

respectively. Note that these entries contain information about the spatially varying coefficient K. The matrix N has the block structure

$$N = \begin{pmatrix} N^x \\ N^y \end{pmatrix},$$

where $N^x \in \mathbb{R}^{(m+1)n \times mn}$ and $N^y \in \mathbb{R}^{m(n+1) \times mn}$ have entries given, respectively, by

$$\int_\Omega \psi_{i,j} \frac{\partial \phi_{k,\ell}^x}{\partial x} dx\, dy, \qquad \int_\Omega \psi_{i,j} \frac{\partial \phi_{k,\ell}^y}{\partial y} dx\, dy.$$

By calculating these integrals, one readily confirms that N^x and N^y reduce to the usual difference approximations to $\partial/\partial x$ and $\partial/\partial y$. The vector $F \in \mathbb{R}^{mn}$ has entries given by the integrals $\int_\Omega f \psi_{i,j} dx\, dy$. The appendix to this paper gives more detail on the construction of A and N.

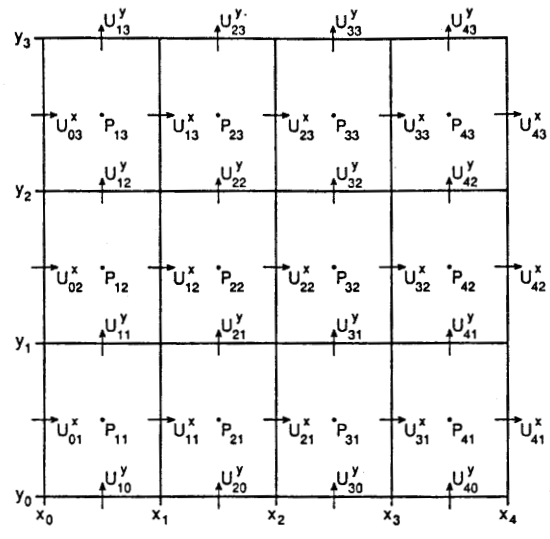

FIG. 1. *Sample* 4×3 *rectangular grid on* $\Omega = (0,1) \times (0,1)$, *showing locations of the nodal unknowns in the velocity and pressure trial functions.*

3. An h-independent iterative method. Our first iterative scheme for solving the discrete system (2.4) is as follows.

ALGORITHM 1. Beginning with initial guess $\left(U^{(0)}, P^{(0)}\right)^T$ for (U, P), the kth iterate $\left(U^{(k)}, P^{(k)}\right)^T$ is the solution of

$$\begin{pmatrix} \omega I & N \\ N^T & 0 \end{pmatrix} \begin{pmatrix} U^{(k)} \\ P^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix} + \begin{pmatrix} \omega I - A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U^{(k-1)} \\ P^{(k-1)} \end{pmatrix}, \tag{3.1}$$

where $I \in \mathbb{R}^{(2mn+m+n)\times(2mn+m+n)}$ is the identity matrix and ω is a parameter chosen to satisfy $\omega \geq \rho(A)$.

Here, $\rho(A)$ denotes the spectral radius of the matrix A. Later in this section we discuss a practical way to pick ω that does not require detailed knowledge of the spectrum of A.

Computationally, Algorithm 1 has the following compact form: Given an initial guess $\left(U^{(0)}, P^{(0)}\right)^T$, compute $\left(U^{(k)}, P^{(k)}\right)^T$ by executing three steps:

$$\text{(3.2)} \quad \text{(i) } G^{(k-1)} \leftarrow -F + \omega^{-1} N^T (\omega I - A) U^{(k-1)},$$

$$\text{(3.3)} \quad \text{(ii) Solve } \omega^{-1} N^T N P^{(k)} = G^{(k-1)},$$

$$\text{(3.4)} \quad \text{(iii) } \omega U^{(k)} \leftarrow (\omega I - A) U^{(k-1)} - N P^{(k)}.$$

In each iteration, the main computational work is to solve for $P^{(k)} = \omega (N^T N)^{-1} G^{(k-1)}$. An easy calculation shows that the matrix $\omega^{-1}(N^T N)$ is positive definite, being pro-

portional to the standard five-point, finite-difference Laplace operator applied to $P^{(k)}$. Therefore, we expect the numerical solution for $P^{(k)}$ using stationary iterative methods to be plagued by poor conditioning when the grid mesh h is small.

This observation leads us to use a multigrid scheme to get approximations to $P^{(k)}$. (In fact, any fast solver for the five-point discrete Laplacian operator would be appropriate here.) Such a device preserves the h-independence of the overall scheme's convergence rate. We discuss this facet of the algorithm in more detail in the next section. For now let us analyze the convergence properties of the overall iterative scheme, assuming an efficient "black-box" solver for $P^{(k)}$.

We begin by writing (3.1) as a stationary iterative scheme

$$\begin{pmatrix} U^{(k)} \\ P^{(k)} \end{pmatrix} = L + M \begin{pmatrix} U^{(k-1)} \\ P^{(k-1)} \end{pmatrix}, \tag{3.5}$$

where

$$L = \begin{pmatrix} \omega I & N \\ N^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ F \end{pmatrix},$$

$$M = \begin{pmatrix} \omega I & N \\ N^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} \omega I - A & 0 \\ 0 & 0 \end{pmatrix}.$$

The convergence of Algorithm 1 depends on the spectral radius of the matrix M, for which the following proposition gives a bound.

PROPOSITION 3.1. *Let*

$$0 < \lambda_{\min} \leq \cdots \leq \lambda_{\max} \tag{3.6}$$

be the eigenvalues of the matrix A, and let $\omega \geq \lambda_{\max}$. Then the spectral radius of M obeys the estimate

$$\rho(M) \leq 1 - \frac{\lambda_{\min}}{\omega}. \tag{3.7}$$

Proof. Let $\lambda \neq 0$ be an eigenvalue of M with eigenvector $(U_\lambda, P_\lambda)^T$. Thus

$$M \begin{pmatrix} U_\lambda \\ P_\lambda \end{pmatrix} \equiv \begin{pmatrix} \omega I & N \\ N^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} \omega I - A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_\lambda \\ P_\lambda \end{pmatrix} = \lambda \begin{pmatrix} U_\lambda \\ P_\lambda \end{pmatrix}, \tag{3.8}$$

so

$$(\omega I - A)U_\lambda = \lambda(\omega U_\lambda + NP_\lambda), \tag{3.9a}$$

$$0 = \lambda N^T U_\lambda. \tag{3.9b}$$

Since $(U_\lambda, P_\lambda)^T \neq 0$, (3.9a) shows that $U_\lambda \neq 0$; however, U_λ may be complex. Let U_λ^H denote its Hermitian conjugate. If we multiply (3.9a) by U_λ^H, observe that N is a real matrix, and apply (3.9b), we obtain

$$U_\lambda^H(\omega I - A)U_\lambda = \lambda\omega U_\lambda^H U_\lambda + \lambda(N^T U_\lambda)^H P_\lambda$$

$$= \lambda\omega U_\lambda^H U_\lambda.$$

This equation allows us to conclude that

$$0 < |\lambda| = \left| \frac{U_\lambda^H (I - \omega^{-1} A) U_\lambda}{U_\lambda^H U_\lambda} \right| \leq \rho(I - \omega^{-1} A),$$

which implies

$$\rho(M) \leq \rho(I - \omega^{-1} A). \tag{3.10}$$

Also, by (3.6) and the fact that $\omega \geq \lambda_{\max}$, we have

$$\rho(I - \omega^{-1} A) \leq 1 - \frac{\lambda_{\min}}{\omega}.$$

These last two inequalities imply the desired bound (3.7). □

If we choose $\omega = \lambda_{\max} = \rho(A)$, then the estimate (3.7) for the spectral radius of the iteration matrix M becomes

$$\rho(M) \leq 1 - \frac{\lambda_{\min}}{\lambda_{\max}}.$$

To estimate $\lambda_{\min}/\lambda_{\max}$, the following proposition is helpful.

PROPOSITION 3.2. *For the matrix A appearing in* (2.4), *there exist constants k_0 and k_1, independent of h, such that*

$$k_0 h^2 U^T U \leq U^T A U \leq k_1 h^2 U^T U. \tag{3.11}$$

Proof. The representation of $\mathbf{u}_h$ given in (2.3) leads to the identity

$$U^T A U = \int_\Omega \frac{1}{K} |\mathbf{u}_h|^2 dx\, dy = \sum_{i=1}^{m} \sum_{j=1}^{n} \int_{\Omega_{i,j}} \frac{1}{K} |\mathbf{u}_h|^2 dx\, dy,$$

where $\Omega_{i,j} = (x_{i-1}, x_i) \times (y_{j-1}, y_j)$. Since K is bounded and integrable on $\Omega_{i,j}$, the mean value theorem for integrals [10, pp. 184–185] guarantees the existence of a number $K_{i,j}$, satisfying $\inf_{\Omega_{i,j}} K \leq K_{i,j} \leq \sup_{\Omega_{i,j}} K$, such that

$$\int_{\Omega_{i,j}} \frac{1}{K} |\mathbf{u}_h|^2 dx\, dy = \frac{1}{K_{i,j}} \int_{\Omega_{i,j}} |\mathbf{u}_h|^2 dx\, dy.$$

(If K^{-1} is continuous on $\bar{\Omega}_{i,j}$, then K^{-1} actually assumes the value $K_{i,j}^{-1}$ somewhere on $\Omega_{i,j}$.) Calculating the last integral using our basis for $\mathbf{Q}_h$, we get

$$U^T A U = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{a_{ij}}{6K_{ij}} \left[\begin{pmatrix} U^x_{i-1,j} \\ U^x_{i,j} \end{pmatrix}^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} U^x_{i-1,j} \\ U^x_{i,j} \end{pmatrix} \right.$$
$$\left. + \begin{pmatrix} U^y_{i,j-1} \\ U^y_{i,j} \end{pmatrix}^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} U^y_{i,j-1} \\ U^y_{i,j} \end{pmatrix} \right],$$

where a_{ij} signifies the area of $\Omega_{i,j}$. To simplify notation, we notice that the 2×2 matrix appearing in each term of this sum is positive definite. This observation allows us to define a new norm on $\mathbb{R}^2$ as follows:

$$\left\| \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \right\|_A^2 = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}.$$

If $\|\cdot\|_2$ denotes the usual Euclidean norm on $\mathbb{R}^2$, then it is easy to check that $\|\cdot\|_2^2 \le \|\cdot\|_A^2 \le 3\|\cdot\|_2^2$. In terms of the new norm,

$$U^T A U = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{a_{ij}}{6K_{ij}} \left[\left\| \begin{pmatrix} U^x_{i-1,j} \\ U^x_{i,j} \end{pmatrix} \right\|_A^2 + \left\| \begin{pmatrix} U^y_{i,j-1} \\ U^y_{i,j} \end{pmatrix} \right\|_A^2 \right].$$

The quantity $U^T U$ is easier to calculate:

$$U^T U = \sum_{i=0}^{m}\sum_{j=1}^{n} |U^x_{ij}|^2 + \sum_{i=1}^{m}\sum_{j=0}^{n} |U^y_{ij}|^2. \tag{3.12}$$

Now we use the bounds on K and the quasi uniformity of Δ_h to observe that

$$\begin{aligned} U^T A U &\ge \frac{\alpha^2 h^2}{6K_{\sup}} \sum_{i=1}^{m}\sum_{j=1}^{n} \left[\left\| \begin{pmatrix} U^x_{i-1,j} \\ U^x_{i,j} \end{pmatrix} \right\|_A^2 + \left\| \begin{pmatrix} U^y_{i,j-1} \\ U^y_{i,j} \end{pmatrix} \right\|_A^2 \right] \\ &\ge \frac{\alpha^2 h^2}{6K_{\sup}} \sum_{i=1}^{m}\sum_{j=1}^{n} \left[\left\| \begin{pmatrix} U^x_{i-1,j} \\ U^x_{i,j} \end{pmatrix} \right\|_2^2 + \left\| \begin{pmatrix} U^y_{i,j-1} \\ U^y_{i,j} \end{pmatrix} \right\|_2^2 \right] \\ &\ge \frac{\alpha^2 h^2}{6K_{\sup}} \left[\sum_{i=0}^{m}\sum_{j=1}^{n} \left\| \begin{pmatrix} 0 \\ U^x_{i,j} \end{pmatrix} \right\|_2^2 + \sum_{i=1}^{m}\sum_{j=0}^{n} \left\| \begin{pmatrix} 0 \\ U^y_{i,j} \end{pmatrix} \right\|_2^2 \right] \\ &= \frac{\alpha^2 h^2}{6K_{\sup}} U^T U. \end{aligned}$$

This observation establishes the first inequality in (3.11), since we can take $k_0 = \alpha^2/6K_{\sup}$. To prove the second inequality in (3.11), we rewrite (3.12) as follows:

$$\begin{aligned} U^T U = \frac{1}{2} &\sum_{i=0}^{m}\sum_{j=0}^{n} \left[\left\| \begin{pmatrix} 0 \\ U^x_{i,j} \end{pmatrix} \right\|_2^2 + \left\| \begin{pmatrix} 0 \\ U^y_{i,j} \end{pmatrix} \right\|_2^2 \right] \\ &+ \frac{1}{2} \sum_{i=1}^{m+1}\sum_{j=1}^{n+1} \left[\left\| \begin{pmatrix} U^x_{i-1,j} \\ 0 \end{pmatrix} \right\|_2^2 + \left\| \begin{pmatrix} U^y_{i,j-1} \\ 0 \end{pmatrix} \right\|_2^2 \right], \end{aligned}$$

where we agree that $U^x_{ij} = 0$ if either $j = 0$ or $j = n+1$, and $U^y_{ij} = 0$ if either $i = 0$ or $i = m+1$. Hence,

$$\begin{aligned} U^T U &\ge \frac{1}{2} \sum_{i=1}^{m}\sum_{j=1}^{n} \left[\left\| \begin{pmatrix} U^x_{i-1,j} \\ U^x_{i,j} \end{pmatrix} \right\|_2^2 + \left\| \begin{pmatrix} U^y_{i,j-1} \\ U^y_{i,j} \end{pmatrix} \right\|_2^2 \right] \\ &\ge \frac{1}{6} \sum_{i=1}^{m}\sum_{j=1}^{n} \left[\left\| \begin{pmatrix} U^x_{i-1,j} \\ U^x_{i,j} \end{pmatrix} \right\|_A^2 + \left\| \begin{pmatrix} U^y_{i,j-1} \\ U^y_{i,j} \end{pmatrix} \right\|_A^2 \right] \\ &\ge \frac{K_{\inf}}{h^2} U^T A U. \end{aligned}$$

We conclude that $U^T A U \le k_1 h^2 U^T U$, where $k_1 = 1/K_{\inf}$. □

If we apply Proposition 3.2 to the case when U is an eigenvector of A associated with the eigenvalue $\lambda_{\min}$ or $\lambda_{\max}$, respectively, we find that $\lambda_{\min} \geq \alpha^2 h^2/6K_{\sup}$ and $\lambda_{\max} \leq h^2/K_{\inf}$. Therefore, provided we choose $\omega \geq \lambda_{\max}$ in Algorithm 1, the spectral radius of our iteration matrix M obeys the bound

$$\rho(M) \leq 1 - \frac{\alpha^2 K_{\inf}}{6K_{\sup}}. \tag{3.13}$$

Notice that the right side of this inequality is a constant independent of h. This is the sense in which the convergence rate of Algorithm 1 is independent of h.

Two remarks about the practical implications of the estimate (3.13) are in order. First, the bound on $\rho(M)$ depends strongly on the nature of the coefficient $K(x, y)$. In particular, if $K_{\inf}/K_{\sup}$ is very small, reflecting a high degree of heterogeneity in the physical problem, then we can expect the actual convergence of the algorithm to be slow, albeit independent of grid mesh. Several examples in §5 confirm this expectation. Second, even though the bound (3.13) suggests choosing $\omega = \lambda_{\max}$ to accelerate iterative convergence, this choice is impractical owing to the expense of calculating $\lambda_{\max}$. In practice, we typically pick $\omega = \|A\|_\infty \geq \lambda_{\max}$. This choice is easily computable as the maximum row sum of A, and it preserves h-independence of convergence rate, even though it may be theoretically nonoptimal.

4. Application of a multigrid solver. As we have mentioned, the computation of the pressure iterate $P^{(k)}$ in step (ii) of Algorithm 1 is inefficient if we use direct schemes or classical stationary iterative methods on fine grids. However, the fact that $\omega^{-1}N^TN$ is essentially the finite-difference Laplacian operator motivates us to reduce the computational work for each iteration by calculating an approximation to the kth pressure iterate by using several cycles of a multigrid method on the system (3.3). We refer the reader to [3] for a discussion of the multigrid approach and for a Fortran code applicable in the context of our problem. The modified scheme is as follows.

ALGORITHM 2. Begin with an initial guess $\left(U^{(0)}, P^{(0)}\right)^T$, and suppose that we have computed $\left(U^{(k-1)}, P^{(k-1)}\right)^T$. Compute a new approximation $\left(U^{(k)}, P^{(k)}\right)^T$ using the following steps:

1. Compute the residual,

$$G^{(k-1)} \leftarrow -F + N^T(I - \omega^{-1}A)U^{(k-1)}. \tag{4.1}$$

2. Let $\hat{P}^{(k)}$ denote the exact solution of the problem

$$\omega^{-1}N^TN\hat{P}^{(k)} = G^{(k-1)}. \tag{4.2}$$

   Calculate an approximation $P^{(k)}$ of $\hat{P}^{(k)}$ by applying r cycles of the multigrid algorithm [3] to (4.2), using $P^{(k-1)}$ as initial guess. (We discuss the choice of r below.)

3. Compute $U^{(k)}$ as in Algorithm 1:

$$\omega U^{(k)} \leftarrow (\omega I - A)U^{(k-1)} - NP^{(k)}. \tag{4.3}$$

Multigrid methods for solving elliptic problems have an advantage that is quite relevant to the conditioning problems associated with fine grids: Each cycle has a convergence rate that is independent of h [4, Chap. 4]. Therefore, we need only show

that we can choose a *fixed* number r of multigrid cycles such that each iteration of Algorithm 2 reduces the error norm by an appropriate factor close to $\rho(M)$. We do this in Proposition 4.1. Since the factor is independent of h, Algorithm 2 has convergence rate independent of h.

We begin by defining norms on the "pressure" and "velocity" spaces that will make the proof easier. Any $p_h \in V_h$ has a representation

$$p_h(x,y) = \sum_{i,j} P_{i,j}\psi_{i,j}(x,y).$$

Taking advantage of the fact that N^TN is positive definite, we compute a norm of the vector

$$P = (P_{1,1}, P_{2,1}, \cdots, P_{m,1}, \cdots, P_{1,n}, P_{2,n}, \cdots, P_{m,n})^T$$

by setting $||P||_h^2 = P^T(\omega^{-1}N^TN)P$. On the other hand, any $\mathbf{u}_h \in \mathbf{Q}_h$ has a representation

$$\mathbf{u}_h(x,y) = \left(\sum_{i,j} U^x_{i,j}\phi^x_{i,j}(x,y), \sum_{i,j} U^y_{i,j}\phi^y_{i,j}(x,y)\right).$$

We compute a norm of the vector

$$U = \big(U^x_{0,1}, U^x_{1,1}, \cdots, U^x_{m,1}, \cdots, U^x_{0,n}, U^x_{1,n}, \cdots, U^x_{m,n},$$

$$U^y_{1,0}, U^y_{1,1}, \cdots, U^y_{1,n}, \cdots, U^y_{m,0}, U^y_{m,1}, \cdots, U^y_{m,n}\big)^T$$

by setting $||U||_\omega^2 = \omega U^TU$.

The norm $\|\cdot\|_\omega$ is just a scalar multiple of the Euclidean distance function $\|\cdot\|_2$, and since ω is a constant related to $\rho(A)$, $\|\cdot\|_\omega$ is actually a discrete analog of the Euclidean norm $\|\cdot\|_{L^2(\Omega)\times L^2(\Omega)}$ on the velocity space by Proposition 3.2. This norm is appropriate for measuring the convergence of velocity iterates $U^{(k)}$ to the true discrete approximation U. Also, since N^TN is just the positive definite matrix associated with the five-point difference approximation to the Laplace operator, the norm $\|\cdot\|_h$ is appropriate for measuring the rapidity with which the pressure iterates satisfy the discrete pressure equation (3.3) as the iterations progress. Ultimately, we want to relate our results to more familiar norms such as $\|\cdot\|_2$ and $\|\cdot\|_\infty$; for this step we shall rely on the equivalence of norms for finite-dimensional Euclidean spaces.

In the following proposition, we assume $\nu = \rho(I-\omega^{-1}A) < 1$. Thus ν is an upper bound on $\rho(M)$. Suppose the multigrid iteration used to approximate $\hat{P}^{(k)}$ in step (ii) of Algorithm 1 has convergence rate $\mu \in (0,1)$. This implies that, after r multigrid cycles for $P^{(k)}$ using $P^{(k-1)}$ as initial guess,

$$\left\|\hat{P}^{(k)} - P^{(k)}\right\|_h \le \mu^r \left\|\hat{P}^{(k)} - P^{(k-1)}\right\|_h. \tag{4.4}$$

PROPOSITION 4.1. *For any $\nu' \in (\nu, 1)$, there exists a number r of multigrid cycles such that*

$$\left\|P - P^{(k)}\right\|_h + \left\|U - U^{(k)}\right\|_\omega \le \nu'\left(\left\|P - P^{(k-1)}\right\|_h + \left\|U - U^{(k-1)}\right\|_\omega\right),$$

where (P,U) is the solution of the problem (2.4) *and* $\left(P^{(k)},U^{(k)}\right)$ *is the approximation to* (P,U) *produced by the kth iteration of Algorithm* 2.

Proof. Suppose we compute $\hat{U}^{(k)}$ according to (3.4) with the exact (nonmultigrid) pressure iterate $\hat{P}^{(k)}$. Thus,

$$\omega\hat{U}^{(k)} = (\omega I - A)U^{(k-1)} - N\hat{P}^{(k)}, \tag{4.5}$$

where $\hat{P}^{(k)}$ satisfies (4.2). Then from (2.4), (4.1), (4.2), and (4.5), we have

$$\omega\left(U-\hat{U}^{(k)}\right) + N\left(P-\hat{P}^{(k)}\right) = (\omega I - A)\left(U-U^{(k-1)}\right), \tag{4.6}$$

$$N^T\left(U-\hat{U}^{(k)}\right) = 0. \tag{4.7}$$

Multiplying (4.6) by $\left(U-\hat{U}^{(k)}\right)^T$ and using the identity (4.7), we get

$$\begin{aligned}\left|\left|U-\hat{U}^{(k)}\right|\right|_\omega^2 &= \left(U-\hat{U}^{(k)}\right)^T(\omega I - A)\left(U-U^{(k-1)}\right)\\ &\le \left|\left|U-\hat{U}^{(k)}\right|\right|_\omega \left|\left|(I-\omega^{-1}A)\left(U-U^{(k-1)}\right)\right|\right|_\omega\\ &\le \rho(I-\omega^{-1}A)\left|\left|U-\hat{U}^{(k)}\right|\right|_\omega \left|\left|U-U^{(k-1)}\right|\right|_\omega.\end{aligned}$$

Therefore, the velocity iterates obey the estimate

$$\left|\left|U-\hat{U}^{(k)}\right|\right|_\omega \le \nu\left|\left|U-U^{(k-1)}\right|\right|_\omega.$$

Similarly, multiplying (4.6) by $\left[\omega^{-1}N\left(P-\hat{P}^{(k)}\right)\right]^T$, we get

$$\begin{aligned}\left|\left|P-\hat{P}^{(k)}\right|\right|_h^2 &= \left(P-\hat{P}^{(k)}\right)^T N^T\omega^{-1}(\omega I - A)\left(U-U^{(k-1)}\right)\\ &\le \left|\left|\omega^{-1}N\left(P-\hat{P}^{(k)}\right)\right|\right|_\omega \left|\left|(I-\omega^{-1}A)\left(U-U^{(k-1)}\right)\right|\right|_\omega\\ &\le \left|\left|P-\hat{P}^{(k)}\right|\right|_h \rho(I-\omega^{-1}A)\left|\left|U-U^{(k-1)}\right|\right|_\omega.\end{aligned}$$

Hence, the pressure iterates obey the bound

$$\left|\left|P-\hat{P}^{(k)}\right|\right|_h \le \nu\left|\left|U-U^{(k-1)}\right|\right|_\omega.$$

Now we derive bounds on $\left|\left|P-P^{(k)}\right|\right|_h$ and $\left|\left|U-U^{(k)}\right|\right|_\omega$ in terms of their values at the previous iterative level. For $\left|\left|P-P^{(k)}\right|\right|_h$, we use the triangle equality and the multigrid estimate (4.4) to get

$$\begin{aligned}\left|\left|P-P^{(k)}\right|\right|_h &\le \left|\left|P-\hat{P}^{(k)}\right|\right|_h + \left|\left|\hat{P}^{(k)}-P^{(k)}\right|\right|_h\\ &\le \left|\left|P-\hat{P}^{(k)}\right|\right|_h + \mu^r\left|\left|\hat{P}^{(k)}-P^{(k-1)}\right|\right|_h \qquad (4.8)\\ &\le \left|\left|P-\hat{P}^{(k)}\right|\right|_h + \mu^r\left(\left|\left|P-\hat{P}^{(k)}\right|\right|_h + \left|\left|P-P^{(k-1)}\right|\right|_h\right).\end{aligned}$$

But the original iterative scheme (3.5) implies that

$$\begin{pmatrix} U - \hat{U}^{(k)} \\ P - \hat{P}^{(k)} \end{pmatrix} = M \begin{pmatrix} U - U^{(k-1)} \\ P - P^{(k-1)} \end{pmatrix}.$$

So, in light of the inequality (3.1) bounding $\rho(M)$ by ν, we have

$$\left|\left|\hat{P} - P^{(k)}\right|\right|_h \leq \rho(M) \left|\left|P - P^{(k-1)}\right|\right|_h \leq \nu \left|\left|P - P^{(k-1)}\right|\right|_h.$$

This inequality allows us to simplify (4.8), getting

$$\left|\left|P - P^{(k)}\right|\right|_h \leq (\nu + \mu^r + \nu\mu^r) \left|\left|P - P^{(k-1)}\right|\right|_h. \tag{4.9}$$

Turning to $||U - U^{(k)}||_\omega$, we use (4.3), multiplied by ω^{-1}, to write

$$\left(U - U^{(k)}\right) = (I - \omega^{-1}A)\left(U - U^{(k-1)}\right) + \omega^{-1}N\left(P - P^{(k)}\right).$$

This identity implies that

$$\begin{aligned} ||U - U^{(k)}||_\omega &\leq ||(I - \omega^{-1}A)\left(U - U^{(k-1)}\right)||_\omega + ||\omega^{-1}N\left(P - P^{(k)}\right)||_\omega \\ &\leq \nu ||U - U^{(k-1)}||_\omega + ||P - P^{(k)}||_h \\ &\leq \nu ||U - U^{(k-1)}||_\omega + (\nu + \mu^r + \nu\mu^r) ||P - P^{(k-1)}||_h \\ &\leq (\nu + \mu^r + \nu\mu^r)\left(||P - P^{(k-1)}||_h + ||U - U^{(k-1)}||_\omega\right). \end{aligned} \tag{4.10}$$

Combining the inequalities (4.9) and (4.10), we get

$$\begin{aligned} &||P - P^{(k)}||_h + ||U - U^{(k)}||_\omega \\ &\quad \leq (\nu + \mu^r + \nu\mu^r)\left(||P - P^{(k-1)}||_h + ||U - U^{(k-1)}||_\omega\right). \end{aligned}$$

Since $\mu < 1$, $\mu^r + \nu\mu^r \to 0$ as $r \to \infty$. We can therefore choose r large enough so that $\nu + \mu^r + \nu\mu^r + \nu \leq \nu' < 1$. In this way,

$$\left|\left|P - P^{(k)}\right|\right|_h + \left|\left|U - U^{(k)}\right|\right|_\omega \leq \nu' \left(\left|\left|P - P^{(k-1)}\right|\right|_h + \left|\left|U - U^{(k-1)}\right|\right|_\omega\right). \qquad \square$$

In view of the norm equivalence mentioned earlier, Proposition 4.1 leads us to expect that, if we choose ω as prescribed in §3, then the computed convergence rate

$$\bar{\mu} = \lim_{k \to \infty} \left[\frac{\|P - P^{(k)}\|_\infty + \|U - U^{(k)}\|_\infty}{\|P - P^{(0)}\|_\infty + \|U - U^{(0)}\|_\infty}\right]^{1/k} \tag{4.11}$$

should be a constant independent of h as $h \to 0$. In fact, for "generic" initial guesses, the contribution from the eigenvector associated with the largest magnitude eigenvalue of M will eventually dominate the error. We therefore expect $\bar{\mu}$ to give good approximations to $\rho(M)$ in computational practice [2, p. 129].

5. Numerical examples of h-independence. To test our results, we apply Algorithm 2 to several versions of the following boundary-value problem:

$$\begin{aligned} -\text{div}\,[K(x,y)\text{grad}\,p(x,y)] &= f(x,y), \qquad (x,y) \in \Omega, \\ p(x,y) &= 0, \qquad (x,y) \in \partial\Omega. \end{aligned} \tag{5.1}$$

We use the lowest-order, mixed finite-element method on grids with $h = 2^{-\ell}$, where $\ell = 4, 5, 6, 7, 8$. Each iteration of the solution scheme includes $r = 2$ V-cycles of the multigrid algorithm described in [3], where the coarsest grid in each cycle has mesh 2^{-1}, and the finest has mesh $2^{-\ell}$. We use the following realizations of the coefficient $K(x,y)$:

$$K_{\text{I}}(x,y) = 1,$$

$$K_{\text{II}}(x,y) = e^{-x-y},$$

$$K_{\text{III}}(x,y) = \begin{cases} 1 & \text{if } x < y, \\ 0.1 & \text{if } x \geq y, \end{cases}$$

$$K_{\text{IV}}(x,y) = K_{\text{II}}(x,y) \cdot K_{\text{III}}(x,y),$$

$$K_{\text{V}}(x,y) = \begin{cases} 1 & \text{if } x < y, \\ 0.01 & \text{if } x \geq y. \end{cases}$$

To confirm the convergence properties of the mixed finite-element method as $h \to 0$, we examine the exact and numerical solutions to (5.1) using $K = K_{\text{II}}$ and taking $f(x,y)$ to be the function that results when the solution is $p(x,y) = x(1-x)\sin(\pi y) + y(1-y)\sin(\pi x)$. We compute the nodal error indicators $\|U_{\text{exact}} - U\|_\infty$ and $\|P_{\text{exact}} - P\|_\infty$, where U_{exact} and P_{exact} stand for the vectors of nodal values of the exact solutions $\mathbf{u}$ and p, and U and P are vectors containing nodal values of the finite-element approximations on a uniform grid of mesh h. Figure 2 shows plots of $\log \|U_{\text{exact}} - U\|_\infty$ and $\log \|P_{\text{exact}} - P\|_\infty$ versus $\log h$ having least-squares slopes of 1.899 and 2.000, respectively. These results suggest that the nodal values of U and P are accurate to $O(h^2)$, corroborating the equal-order accuracy available in the Raviart–Thomas subspaces and indicating superconvergent nodal values in accordance with the work of Ewing, Lazarov, and Wang [6].

To check the convergence properties of the iterative scheme, we examine the behavior of the ratio $\bar{\mu}$, defined in (4.11), for each of the choices of K. Our results, shown in Fig. 3, support the expectation that, as $h \to 0$, the convergence rate of the scheme tends to a constant independent of h. Notice however that, as K exhibits more spatial variation, the convergence of the algorithm becomes slower. Any effects of variability in K on the conditioning of the discrete equations still influence this first algorithm; the only effects of poor conditioning that we have eliminated so far are those associated with grid refinement.

6. Modified schemes for heterogeneous media. To mitigate the difficulties associated with spatial variability, we modify the first iterative scheme (3.1) to get a class of new schemes having the following form.

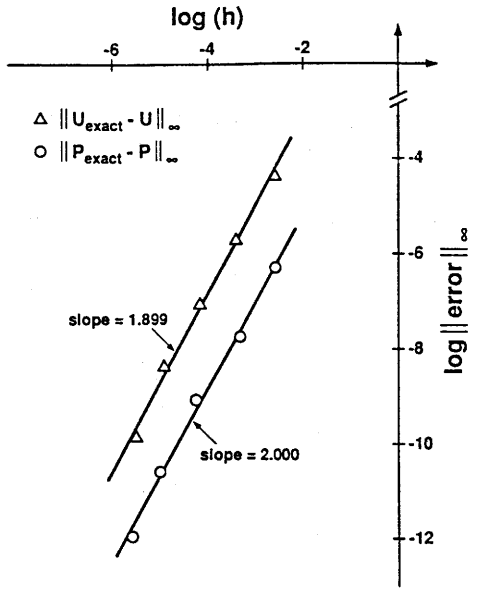

FIG. 2. *Convergence plot for the mixed finite-element scheme for Poisson's equation, using lowest-order Raviart–Thomas trial spaces. The plots demonstrate the rate of decrease in the nodal errors as* $h \to 0$.

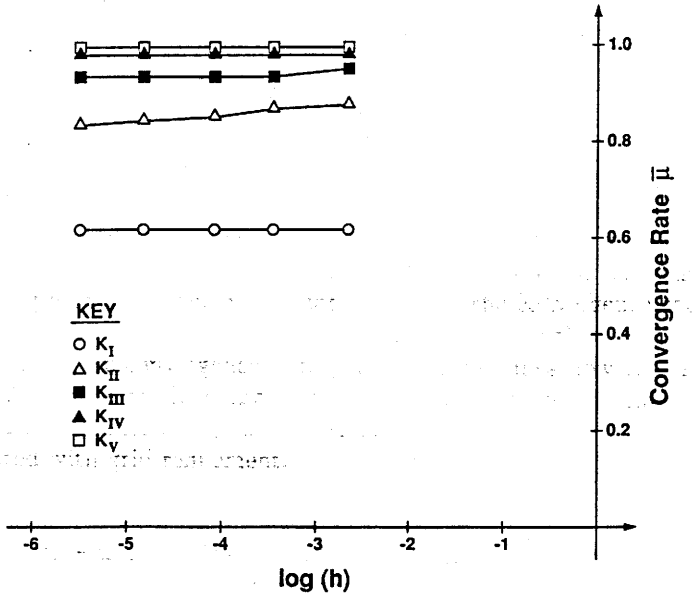

FIG. 3. *Rate of convergence* $\bar{\mu}$ *versus grid mesh* h *for Algorithm 2, using the various choices of coefficient* $K(x, y)$.

ALGORITHM 3. Given initial guess $\left(U^{(0)}, P^{(0)}\right)^T$, the kth iterate $\left(U^{(k)}, P^{(k)}\right)^T$ is the solution of

$$\begin{pmatrix} D & N \\ N^T & 0 \end{pmatrix} \begin{pmatrix} U^{(k)} \\ P^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix} + \begin{pmatrix} D-A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U^{(k-1)} \\ P^{(k-1)} \end{pmatrix}. \tag{6.1}$$

Here, the "preconditioning" matrix $D \in \mathbb{R}^{(2mn+m+n)\times(2mn+m+n)}$ is a diagonal matrix whose choice we discuss below.

When we construct D properly, the iteration matrix

$$M = \begin{pmatrix} D & N \\ N^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} D-A & 0 \\ 0 & 0 \end{pmatrix} \tag{6.2}$$

has spectral radius that is independent of both h and the structure of K. The price we pay for this benefit is apparent in the computational form of the new algorithm:

$$\text{(i) } G^{(k-1)} \leftarrow -F + N^T D^{-1}(D-A)U^{(k-1)}, \tag{6.3}$$

$$\text{(ii) Solve } N^T D^{-1} N P^{(k)} = G^{(k-1)}, \tag{6.4}$$

$$\text{(iii) } U^{(k)} \leftarrow D^{-1}(D-A)U^{(k-1)} - D^{-1}NP^{(k)}. \tag{6.5}$$

In contrast to (3.3), solving for $P^{(k)}$ in the new scheme calls for the inversion of $N^T D^{-1} N$ instead of $N^T N$. Therefore, we must modify the multigrid segment of the algorithm to accommodate variable coefficients. As we discuss, this modification is fairly easy to make. This section establishes criteria for the construction of D, gives two examples that satisfy these criteria, comments on the multigrid solver used, and presents computational results.

As with the original scheme presented in §3, the key to the convergence of the new scheme is the spectral radius of the iteration matrix M defined in (6.2). The following proposition gives sufficient conditions under which $\rho(M) < 1$.

PROPOSITION 6.1. *Suppose D is a diagonal matrix with positive entries on the diagonal, and suppose there exist constants $b_1, b_2 \in (0,1)$ such that*

$$b_1 \le \frac{U^H A U}{U^H D U} \le 2 - b_2$$

for all vectors $U \in \mathbb{C}^{(m+1)n+m(n+1)}$. Then the iteration matrix M defined in (6.2) *satisfies*

$$0 < \rho(M) \le \max\{1-b_1, 1-b_2\} < 1. \tag{6.6}$$

Proof. Let $\lambda \neq 0$ be an eigenvalue of M with associated eigenvector $(U_\lambda, P_\lambda)^T$, as in Proposition 3.1. Then steps similar to those yielding (3.9) show that

$$\begin{aligned} (D-A)U_\lambda &= \lambda(DU_\lambda + NP_\lambda), \\ 0 &= \lambda N^T U_\lambda. \end{aligned}$$

Thus $U_\lambda^H (D-A)U_\lambda = \lambda U_\lambda^H D U_\lambda$, which is nonzero since D is positive definite. Therefore,

$$|\lambda| = \left| 1 - \frac{U_\lambda^H A U_\lambda}{U_\lambda^H D U_\lambda} \right|.$$

Hence, using the hypothesized bounds on $U_\lambda^H AU_\lambda / U_\lambda^H DU_\lambda$, we have the desired inequalities (6.6). □

To use this proposition, we need estimates on $U^H AU$. Given the structure of A as shown in the Appendix, one can calculate a useful expression for $U^H AU$, assuming $U \in \mathbb{C}^{(m+1)n+m(n+1)}$ has the form $(U^x, U^y)^T$ indicated in (2.3). In particular,

$$U^H AU = \frac{1}{3}S(U) + \frac{1}{6}R(U),$$

where, in the notation of the Appendix,

$$S(U) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(T_{i,j}^{\mathrm{I}}|U_{i-1,j}^x|^2 + T_{i,j}^{\mathrm{III}}|U_{i,j}^x|^2 + T_{i,j}^{\mathrm{IV}}|U_{i,j-1}^y|^2 + T_{i,j}^{\mathrm{VI}}|U_{i,j}^y|^2\right),$$

$$R(U) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left[T_{i,j}^{\mathrm{II}}\left(\bar{U}_{i,j}^x U_{i-1,j}^x + \bar{U}_{i-1,j}^x U_{i,j}^x\right) + T_{i,j}^{\mathrm{V}}\left(\bar{U}_{i,j}^y U_{i,j-1}^y + \bar{U}_{i,j-1}^y U_{i,j}^y\right)\right].$$

Here, $\bar{z}$ denotes the complex conjugate of z. The coefficients $T_{i,j}^{\mathrm{I}}, \cdots, T_{i,j}^{\mathrm{VI}}$ appearing in these expressions are values depending on $K(x,y)$ and arising from applications of the mean value theorem for integrals over each cell $\bar{\Omega}_{i,j}$ in the finite-element grid Δ_h. By using the inequality $|w|^2 + |z|^2 \geq -2|w||z|$, we can estimate $R(U)$ as follows:

(6.7)

$$-2\sum_{i=1}^{m}\sum_{j=1}^{n}\left(T_{i,j}^{\mathrm{II}}|U_{i,j}^x||U_{i-1,j}^x| + T_{i,j}^{\mathrm{V}}|U_{i,j}^y||U_{i,j-1}^y|\right) \leq R(U)$$

$$\leq \sum_{i=1}^{m}\sum_{j=1}^{n}\left(T_{i,j}^{\mathrm{II}}|U_{i-1,j}^x|^2 + T_{i,j}^{\mathrm{II}}|U_{i,j}^x|^2 + T_{i,j}^{\mathrm{V}}|U_{i,j-1}^y|^2 + T_{i,j}^{\mathrm{V}}|U_{i,j}^y|^2\right).$$

In general, the estimates $0 < K_{\mathrm{inf}} \leq K \leq K_{\mathrm{sup}}$ may be too coarse to provide enough control on the coefficients $T_{i,j}^{\mathrm{I}}, \cdots, T_{i,j}^{\mathrm{VI}}$ for constructing a reasonable preconditioner D. Strictly speaking, the necessary level of control will be available only if we have information about the *local* variation of K on each cell $\bar{\Omega}_{i,j}$.

In practice, however, we rarely have such fine-scale knowledge of K, and even if we did we would not try to use it in calculating the Galerkin integrals $\int_\Omega K^{-1}\mathbf{u}\cdot\mathbf{v}\,dx\,dy$ exactly. Instead, most practical codes use approximate quadrature schemes that effectively treat K^{-1} as piecewise polynomial. In fact, as we suggested in §1, for sufficiently fine grids it is reasonable to treat K^{-1} as piecewise constant. In such applications, we can use the second inequality in (6.7), together with the identities $T_{i,j}^{\mathrm{II}} = T_{i,j}^{\mathrm{V}} = T_{i,j}$, to show that

$$U^H AU = \frac{1}{3}S(U) + \frac{1}{6}R(U) \leq \frac{1}{2}S(U).$$

Similarly, the first inequality in (6.7), together with the identities $T_{i,j}^{\mathrm{I}} = T_{i,j}^{\mathrm{III}} = T_{i,j}^{\mathrm{IV}} =$

$T_{i,j}^{\mathrm{VI}} = T_{i,j}$, shows that

$$\begin{aligned} U^H A U &= \frac{1}{6}S(U) + \frac{1}{6}[S(U) + R(U)] \\ &\geq \frac{1}{6}S(U) + \frac{1}{6}\sum_{i=1}^{m}\sum_{j=1}^{n} T_{i,j}\left[\left(|U_{i-1,j}^x| - |U_{i,j}^x|\right)^2 + \left(|U_{i,j-1}^y| - |U_{i,j}^y|\right)^2\right] \\ &\geq \frac{1}{6}S(U). \end{aligned}$$

In summary, $\frac{1}{6}S(U) \leq U^H A U \leq \frac{1}{2}S(U)$ whenever K is piecewise constant on the grid Δ_h.

Now consider the choice $D = \frac{2}{3}\mathrm{lump}(A)$, where

$$[\mathrm{lump}(A)]_{i,j} = \begin{cases} 0 & \text{if } i \neq j, \\ \displaystyle\sum_j A_{i,j} & \text{if } i = j. \end{cases}$$

This is the matrix that results when we add entries along each row of A and assign the sum to the diagonal entry in that row. Gonzales and Wheeler [9] use this "mass lumping" idea to improve conditioning in mixed finite-element discretizations of petroleum reservoir problems. This choice of D is also a simple instance of a preconditioner developed in [7] for other iterative schemes. It is a straightforward matter to show that, when K is piecewise constant, $U^H \mathrm{lump}(A) U = \frac{1}{2}S(U)$, so $U^H D U = \frac{1}{3}S(U)$. As a consequence,

$$b_1 = \frac{1}{2} \leq \frac{U^H A U}{U^H D U} \leq \frac{3}{2} = 2 - b_2.$$

Therefore, by Proposition 6.1, $\rho(M) \leq \frac{1}{2}$, and the iterative scheme converges with a rate independent of h and K. According to our remarks at the end of §4, we expect the ratio of error norms between successive iterates to approach $\frac{1}{2}$ as the iteration counter $k \to \infty$.

As an even simpler example, consider the choice $D = \mathrm{diag}(A)$, where

$$[\mathrm{diag}(A)]_{i,j} = \begin{cases} 0 & \text{if } i \neq j, \\ A_{i,i} & \text{if } i = j, \end{cases}$$

is the matrix A stripped of its off-diagonal entries. This choice has the attractive feature that it is trivial to compute from A. With D defined in this way, we once again find that $U^H D U = \frac{1}{3}S(U)$ when K is piecewise constant on Δ_h. Therefore, $\rho(M) \leq \frac{1}{2}$, and this iterative scheme also converges with a rate independent of h and K.

Either choice of D requires us to solve a matrix equation of the form

$$N^T D^{-1} N P^{(k)} = G^{(k-1)}$$

at each iteration. To do this, we use two cycles of a multigrid scheme in which the Jacobi iteration is the smoother, the coarse-to-fine interpolation is bilinear, and the fine-to-coarse restriction is accomplished using half-injection [4, p. 65]. This scheme preserves the h-independence of the overall algorithm's convergence rate and appears

TABLE 1
Convergence rates for various coefficients and grids.

	Grid Mesh h				
Coefficient	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}
K_{I}	0.4933	0.4988	0.4993	0.4995	0.4999
K_{II}	0.4966	0.4995	0.4988	0.4997	0.4999
K_{III}	0.4948	0.4982	0.4991	0.4998	0.4999
K_{IV}	0.4947	0.4980	0.4992	0.4998	0.4999
K_{V}	0.4939	0.4978	0.4989	0.4999	0.5000

to handle the variable coefficient K effectively. Alternative multigrid implementations are certainly possible here.

To test the convergence rate of Algorithm 3, we apply it to the boundary-value problems described in §5, using the preconditioner $D = \frac{2}{3}\mathrm{lump}(A)$. Table 1 shows values of the convergence rate $\bar{\mu}$ computed for each choice of coefficient K, for each of five different values of the grid mesh h. All of the tabulated values are very close to the spectral radius estimate $\rho(M) \leq \frac{1}{2}$. We conclude that this scheme converges at a rate independent of both grid mesh h and the heterogeneity reflected in the mobility coefficient K.

7. Conclusions. Poor conditioning associated with heterogeneity and fine spatial grids is a common problem. While this paper focuses on steady flows in porous media, similar equations and results apply in other fields. Two obvious applications for (1.1) arise in heat transfer, where temperature plays the role of pressure and heat flux plays the role of the Darcy velocity, and in electrostatics, where the electric potential and the electric field serve as the analogs of pressure and Darcy velocity, respectively. In either case, mixed finite-element methods can give useful approximations. However, heterogeneity, either in the thermal diffusivity or in the dielectric coefficient, can lead to poor conditioning in precisely the same way as it does for porous media. One virtue of the mixed finite-element formulation is that it permits us to attack the two sources of poor conditioning separately, exploiting multigrid ideas to reduce the sensitivity to fine grids and using spectral information associated with the material coefficient to reduce the sensitivity to heterogeneity.

Appendix: Matrix structure of the finite-element equations. The mixed finite-element equations (2.2) give rise to integral equations having the following forms. For the x-velocity equation,

$$\int_\Omega \left(K^{-1} u_h^x \, \phi_{i,j}^x - p_h \frac{\partial \phi_{i,j}^x}{\partial x} \right) dx\, dy = 0, \quad i = 0, \cdots, m, \quad j = 1, \cdots, n.$$

For the y-velocity equation,

$$\int_\Omega \left(K^{-1} u_h^y \, \phi_{i,j}^y - p_h \frac{\partial \phi_{i,j}^y}{\partial y} \right) dx\, dy = 0, \quad i = 1, \cdots, m, \quad j = 0, \cdots, n.$$

For the mass balance,

$$\int_\Omega \left(\frac{\partial u_h^x}{\partial x} + \frac{\partial u_h^y}{\partial y} - f \right) \psi_{i,j} dx\, dy = 0, \quad i = 1, \cdots, m, \quad j = 1, \cdots, n.$$

The following integrals appearing in these expressions involve no spatially varying coefficients and hence are easy to compute using the bases for $\mathbf{Q}_h$ and V_h:

$$\int_\Omega p_h \frac{\partial \phi_{i,j}^x}{\partial x}\,dx\,dy, \quad \int_\Omega p_h \frac{\partial \phi_{i,j}^y}{\partial y}\,dx\,dy, \quad \int_\Omega \frac{\partial u_h^x}{\partial x}\psi_{i,j}\,dx\,dy, \quad \int_\Omega \frac{\partial u_h^y}{\partial y}\psi_{i,j}\,dx\,dy.$$

However, the remaining integrals involve the spatially varying functions $K^{-1}(x,y)$ and $f(x,y)$. We compute these integrals using the mean value theorem for integrals [10, pp. 184–185] as follows: Since K^{-1} is bounded and integrable on each cell $\overline{\Omega}_{i,j}$, there exist numbers $T_{i,j}^{\mathrm{I}}, T_{i,j}^{\mathrm{II}}, T_{i,j}^{\mathrm{III}}$ such that

$$\int_\Omega K^{-1}\,\phi_{s,t}^x\,\phi_{i,j}^x\,dx\,dy = \begin{cases} T_{i,j}^{\mathrm{II}}/6, & t=j, \quad s=i-1; \\ \left(T_{i,j}^{\mathrm{I}} + T_{i+1,j}^{\mathrm{III}}\right)/3, & t=j, \quad s=i; \\ T_{i+1,j}^{\mathrm{II}}/6, & t=j, \quad s=i+1\,. \end{cases}$$

Here, $T_{i,j}^{\mathrm{I}}/[(x_i - x_{i-1})(y_j - y_{j-1})]$ is a number lying between the upper and lower bounds of K^{-1} on the cell $\overline{\Omega}_{i,j}$, and similarly for $T_{i,j}^{\mathrm{II}}$ and $T_{i,j}^{\mathrm{III}}$. Analogous calculations show that

$$\int_\Omega K^{-1}\,\phi_{s,t}^y\,\phi_{i,j}^y\,dx\,dy = \begin{cases} T_{i,j}^{\mathrm{V}}/6, & t=j-1, \quad s=i; \\ \left(T_{i,j}^{\mathrm{IV}} + T_{i,j+1}^{\mathrm{VI}}\right)/3, & t=j, \quad s=i; \\ T_{i,j+1}^{\mathrm{V}}/6, & t=j+1, \quad s=i\,. \end{cases}$$

The calculations of $\int_\Omega f\,\psi_{i,j}\,dx\,dy$ can proceed similarly.

Now let us adopt the following orderings for the vectors of unknown nodal coefficients:

$$U^x = \begin{bmatrix} U_{0,1}^x \\ \vdots \\ U_{m,1}^x \\ \vdots \\ U_{0,n}^x \\ \vdots \\ U_{m,n}^x \end{bmatrix}, \quad U^y = \begin{bmatrix} U_{1,0}^y \\ \vdots \\ U_{1,n}^y \\ \vdots \\ U_{m,0}^y \\ \vdots \\ U_{m,n}^y \end{bmatrix}, \quad P = \begin{bmatrix} P_{1,1} \\ \vdots \\ P_{m,1} \\ \vdots \\ P_{1,n} \\ \vdots \\ P_{m,n} \end{bmatrix}.$$

Then the entire algebraic system arising from (2.2) has the structure

$$\begin{bmatrix} A^x & 0 & N^x \\ 0 & A^y & N^y \\ (N^x)^T & (N^y)^T & 0 \end{bmatrix} \begin{bmatrix} U^x \\ U^y \\ P \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ F \end{bmatrix}.$$

Here,

$$A^x = \begin{bmatrix} A_1^x & & \\ & \ddots & \\ & & A_n^x \end{bmatrix} \in \mathbb{R}^{(m+1)n \times (m+1)n},$$

where each block $A_j^x \in \mathbb{R}^{(m+1)\times(m+1)}$ has the tridiagonal structure

$$A_j^x = \frac{1}{6}\begin{bmatrix} 2T_{1,j}^{\mathrm{III}} & T_{1,j}^{\mathrm{II}} & & & \\ T_{1,j}^{\mathrm{II}} & 2(T_{1,j}^{\mathrm{III}} + T_{2,j}^{\mathrm{I}}) & T_{2,j}^{\mathrm{II}} & & \\ & & \ddots & & \\ & & & T_{m,j}^{\mathrm{II}} & 2T_{m,j}^{\mathrm{III}} \end{bmatrix}.$$

Similarly,

$$A^y = \begin{bmatrix} A_1^y & & \\ & \ddots & \\ & & A_m^y \end{bmatrix} \in \mathbb{R}^{m(n+1)\times m(n+1)},$$

where each block $A_i^y \in \mathbb{R}^{(n+1)\times(n+1)}$ has the tridiagonal form

$$A_i^y = \frac{1}{6}\begin{bmatrix} 2T_{i,1}^{\mathrm{VI}} & T_{i,1}^{\mathrm{V}} & & & \\ T_{i,1}^{\mathrm{V}} & 2(T_{i,1}^{\mathrm{VI}} + T_{i,2}^{\mathrm{IV}}) & T_{i,2}^{\mathrm{V}} & & \\ & & \ddots & & \\ & & & T_{i,n}^{\mathrm{V}} & 2T_{i,n}^{\mathrm{VI}} \end{bmatrix}.$$

Finally, the two "difference" matrices N^x and N^y have the following structures:

$$N^x = \begin{bmatrix} N_1^x & & \\ & \ddots & \\ & & N_n^x \end{bmatrix} \in \mathbb{R}^{n(m+1)\times nm},$$

where

$$N_j^x = (y_j - y_{j-1})\begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & -1 \end{bmatrix} \in \mathbb{R}^{(m+1)\times m},$$

while

$$N^y = \begin{bmatrix} N_{1,1}^y & \cdots & N_{1,n}^y \\ \vdots & & \\ N_{m,1}^y & \cdots & N_{m,n}^y \end{bmatrix} \in \mathbb{R}^{(n+1)m\times nm},$$

where

$$N_{i,j}^{y} = (x_i - x_{i-1}) \begin{bmatrix} & \vdots & \\ \cdots & 1 & \cdots \\ \cdots & -1 & \cdots \\ & \vdots & \end{bmatrix} \begin{matrix} \\ \leftarrow \text{ row } j \\ \leftarrow \text{ row } j+1 \\ \\ \end{matrix} \in \mathbb{R}^{(n+1)\times m}.$$

$$\uparrow$$
$$\text{column } i$$

REFERENCES

[1] M. B. ALLEN, R. E. EWING, AND J. V. KOEBBE, *Mixed finite-element methods for computing groundwater velocities*, Numer. Meth. P.D.E., 3 (1985), pp. 195–207.

[2] G. BIRKHOFF AND R. E. LYNCH, *Numerical Solution of Elliptic Problems*, Society for Industrial and Applied Mathematics, Philadelphia, 1984.

[3] A. BRANDT, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[4] W. L. BRIGGS, *A Multigrid Tutorial*, Society for Industrial and Applied Mathematics, Philadelphia, 1987.

[5] J. DOUGLAS, R. E. EWING, AND M. F. WHEELER, *The approximation of the pressure by a mixed method in the simulation of miscible displacement*, RAIRO Anal. Numér., 17 (1983), pp. 17–33.

[6] R. E. EWING, R. D. LAZAROV, AND J. WANG, *Superconvergence of the velocities along the Gaussian lines in mixed finite element methods*, SIAM J. Numer. Anal., 28 (1991), pp. 1015–1029.

[7] R. E. EWING, R. D. LAZAROV, P. LU, AND P. S. VASSILEVSKI, *Preconditioning indefinite systems arising from the mixed finite-element discretization of second-order elliptic systems*, in Preconditioned Conjugate Gradient Methods, Lecture Notes in Mathematics 1457, O. Axelsson and L. Kolotilina, eds., Springer-Verlag, Berlin, 1990, pp. 280–343.

[8] R. E. EWING AND M. F. WHEELER, *Computational aspects of mixed finite element methods*, in Numerical Methods for Scientific Computing, R. S. Stepelman, ed., North-Holland, Amsterdam, 1983, pp. 163–172.

[9] R. GONZALES AND M. F. WHEELER, *Mixed finite element methods for petroleum reservoir engineering problems*, in Proceedings, Sixth International Conference on Computing Methods in Applied Sciences and Engineering, INRIA, Versailles, France, 1983, North-Holland, Amsterdam, 1984, pp. 639–658.

[10] M. E. MUNROE, *Introduction to Measure and Integration*, Addison-Wesley, Cambridge, MA, 1953.

[11] P. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical Aspects of Finite Element Methods, Lecture Notes in Mathematics 606, I. Galligani and E. Magenes, eds., Springer-Verlag, Berlin, 1977, pp. 292–315.