SIAM J. SCI. COMPUT.
Vol. 17, No. 1, pp. 33–46, January 1996

003

FAST NONSYMMETRIC ITERATIONS AND PRECONDITIONING FOR NAVIER–STOKES EQUATIONS*

HOWARD ELMAN[†] AND DAVID SILVESTER[‡]

Abstract. Discretization and linearization of the steady-state Navier–Stokes equations gives rise to a nonsymmetric indefinite linear system of equations. In this paper, we introduce preconditioning techniques for such systems with the property that the eigenvalues of the preconditioned matrices are bounded independently of the mesh size used in the discretization. We confirm and supplement these analytic results with a series of numerical experiments indicating that Krylov subspace iterative methods for nonsymmetric systems display rates of convergence that are independent of the mesh parameter. In addition, we show that preconditioning costs can be kept small by using iterative methods for some intermediate steps performed by the preconditioner.

Key words. Navier–Stokes, iterative methods, preconditioners, Krylov subspace

AMS subject classifications. 65F10, 65N12, 65N22, 65M60

1. Introduction. Consider the steady-state *Navier–Stokes problem*: given data $\mathbf{f}$, find the velocity $\mathbf{u}$ and pressure p satisfying

$$\begin{aligned} -\nu\nabla^2\mathbf{u} + \frac{1}{2}\mathbf{u}(\operatorname{div}\mathbf{u}) + \mathbf{u}\cdot\nabla\mathbf{u} + \operatorname{grad} p &= \mathbf{f} \\ \operatorname{div}\mathbf{u} &= 0 \end{aligned} \qquad \text{in } \Omega \tag{1.1}$$

subject to boundary conditions, typically, specified velocity $\mathbf{u} = \mathbf{g}$ on $\partial\Omega$; $\Omega \subset \mathbf{R}^2$ or $\Omega \subset \mathbf{R}^3$. Here, the scalar ν is the inverse of the Reynolds number or the ratio of convection to diffusion in the system. In the diffusion dominated case, ($\nu \to \infty$) (1.1) tends to a linear self-adjoint system of equations known as the *Stokes problem*.

There are two ways of calculating solutions to the system (1.1). A popular approach is to compute "true" steady-state solutions of the time-dependent Navier–Stokes equations. There are many ways to do this: one way is to make use of the "characteristics" associated with the hyperbolic part of the Navier–Stokes operator via a Lagrange–Galerkin approach (for example, see [13]). The associated transpose-diffusion splitting leads to absolutely stable temporal discretizations so that large time steps can be taken. At each time step, a symmetric indefinite matrix system corresponding to a time-discretized Stokes-like system must be solved. These systems can be solved efficiently by iterative methods, for example, if a multigrid solver is used to precondition the primary (Laplacian) operator. There are, however, a number of disadvantages to the time-dependent approach. Simple time discretization methods based on the l_2-projection onto the discretely divergence-free subspace [10] have an $O(h)$ CFL restriction on the time step, which impinges on efficiency. On the other hand, absolutely stable schemes like the method of backward characteristics are known to be sensitive to implementation issues (e.g., the need to perform quadrature; see [13]). Even with fixed grids, efficiency is often limited by the costs associated with interpolation.

In this work, we consider the alternative approach of attacking the system (1.1) directly. Applying a fixed point (or Picard) iteration, the system (1.1) reduces to solving a sequence of

*Received by the editors June 8, 1994; accepted for publication (in revised form) January 21, 1995. This work was supported by the U. S. Army Research Office under grant DAAL-0392-G-0016 and the U. S. National Science Foundation under grant ASC-8958544 at the University of Maryland and the Science and Engineering Research Council of Great Britain Visiting Fellowship GR/J 04982 at the University of Manchester Institute of Science and Technology.

[†] Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 (elman@cs.umd.edu).

[‡] Department of Mathematics, University of Manchester Institute of Science and Technology, Manchester M601QD, UK (na.silvester@na-net.ornl.gov).

linear Oseen problems of the following form: given some velocity field $\mathbf{w}$, find the velocity $\mathbf{u}$ and pressure p satisfying

$$(1.2)\qquad \begin{aligned} -\nu\nabla^2\mathbf{u} + \frac{1}{2}\mathbf{u}(\operatorname{div}\mathbf{w}) + \mathbf{w}\cdot\nabla\mathbf{u} + \operatorname{grad} p &= \mathbf{f} \\ \operatorname{div}\mathbf{u} &= 0 \end{aligned} \qquad \text{in } \Omega$$

subject to the same boundary conditions.

For this methodology to be effective it is necessary to solve the discrete versions of (1.2) efficiently. Finite element discretization of (1.2) leads to the matrix problem

$$(1.3)\qquad \begin{pmatrix} \nu A + N & B^t \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

where $A = A^t$ represents diffusion (for example, $-\nabla^2$) and hence is a positive definite matrix of order n_u; N represents convection ($\frac{1}{2}\operatorname{div}\mathbf{w} + \mathbf{w}\cdot\nabla$); and the $n_p \times n_u$ matrix B represents the coupling between the discrete velocity u and the pressure p. Note that the representation of the quadratic convection term in (1.1) ensures that if velocity is specified everywhere on the boundary then $N = -N^t$; that is, the discrete form of the convection operator is *skew-symmetric* [10, p. 53]. Note that with these boundary conditions the system (1.3) is singular; pressure is only unique up to a (hydrostatic) constant. We assume in the following analysis that the pressure solution is uniquely specified in this case, e.g., by insisting that its mean is zero.

Working in a conventional mixed finite element framework, we will further assume that the underlying velocity and pressure approximations are *(div-)stable* (see, e.g., [2, p. 57], [10, pp. 10ff], and [19]); i.e., defining a mesh parameter h, a velocity space $\mathbf{V}_h$, and a pressure space P_h, there exist constants γ, Γ, independent of h, such that

$$(1.4)\qquad \gamma^2 \le \frac{(p, BA^{-1}B^t p)}{(p, Qp)} \le \Gamma^2 \qquad \forall p \in P_h .$$

Here, Q is the pressure mass matrix, or alternatively the Grammian matrix of basis functions defining P_h. The lower bound γ is the so-called *inf–sup* constant. The relation (1.4) is crucial to the success of iterative solvers for solving discrete Stokes problems because it implies that, using a quasi-uniform mesh, the Schur complement $BA^{-1}B^t$ has a condition number bounded independently of h. It is also known from our previous work [16] that when $\nu \to \infty$, "optimal" preconditioners for the Laplacian subblocks give rise to "optimal" preconditioners for the Stokes problem in the sense that the spectra of the underlying discrete operators are contained in small clusters, which are bounded independently of h. A consequence of this is that the asymptotic convergence rate, with respect to the preconditioned residual norm, of Krylov subspace methods applied to discrete Stokes problems is also independent of h.

In this paper, we derive analogous results for eigenvalues in the general Oseen case. We introduce two preconditioners for the Oseen problem such that, for any value $0 < \nu < \infty$, the eigenvalues of the preconditioned Oseen operator are bounded independently of the mesh size. These observations apply to arbitrary discretizations satisfying (1.4). In addition, in a series of numerical experiments we show that these bounds on eigenvalues are predictive of the performance of Krylov subspace iterative methods for solving the preconditioned Oseen equations. Of course, it is well known that when convection dominates (i.e., when ν is "small" relative to h and $\|\mathbf{w}\|$), the standard Galerkin approximation deteriorates. Oscillations in the discrete velocity are apparent if the local mesh Reynolds number $Re^h = h\|\mathbf{w}\|/\nu$ is greater than unity. In such situations, the addition of streamwise diffusion to the discrete system is known to give added stability, both theoretically and numerically; see [3] and [12]. In our experiments, we demonstrate the effectiveness of the ideas using both a standard Galerkin discretization on a set of quasi-uniform grids and a streamline-upwind scheme on a set of uniform grids.

The remainder of the paper is divided into three sections. Our main theoretical results are presented in §2, and results of numerical experiments confirming and augmenting the theoretical analysis are given in §3. In §4, we consider more practical preconditioning strategies and present a perturbation analysis and additional numerical experiments demonstrating their effectiveness.

2. Preconditioning strategies. In this section, we introduce two preconditioning techniques for (1.1) and present an analysis showing that the spectra of the preconditioned systems are bounded independently of the discretization mesh size h. Throughout the section, we will be concerned with the eigenvalues of preconditioned matrices; these matrices can be viewed as being of the form $\mathcal{A}\mathcal{M}^{-1}$, where $\mathcal{A}$ is the original matrix and $\mathcal{M}$ is the preconditioner. Equivalently, we are concerned with the solution of the generalized eigenvalue problem $\mathcal{A}v = \lambda\mathcal{M}v$. All the matrices in question are implicitly parameterized by h. For simplicity, we state our results under the assumption that B of (1.3) has full rank.

The first idea is derived from a method developed in [14], [16], [20] for the discrete Stokes equations, where the coefficient matrix has the form

$$\begin{pmatrix} A & B^t \\ B & O \end{pmatrix}. \tag{2.1}$$

Consider the preconditioner

$$\begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix}$$

for (2.1). The eigenvalues of the preconditioned operator are then given by the solution to the generalized eigenvalue problem

$$\begin{pmatrix} A & B^t \\ B & O \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix} = \lambda \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix}. \tag{2.2}$$

One solution is $\lambda = 1$, of multiplicity $n_u - n_p$, for which the eigenvectors have the form $\binom{u}{0}$ where $Bu = 0$; i.e., u is "discretely divergence free." The remaining eigenvalues come from the solution of the quadratic equation $\lambda(\lambda - 1) = \mu$, where μ is a generalized eigenvalue of the Schur complement associated with (2.1),

$$BA^{-1}B^t p = \mu\, Qp\,. \tag{2.3}$$

Equivalently,

$$\lambda = \frac{1 \pm \sqrt{1 + 4\mu}}{2}\,. \tag{2.4}$$

Since (1.4) implies that as $h \to 0$ the solutions to (2.3) remain bounded above and below, it follows that the eigenvalues of (2.2) are also bounded. The preconditioned conjugate residual method can then be used to solve (2.1), with a convergence rate independent of h [14], [16].

A natural generalization for the discrete Oseen equations uses the block preconditioner

$$\begin{pmatrix} F & 0 \\ 0 & \frac{1}{\nu}Q \end{pmatrix}, \tag{2.5}$$

where $F = \nu A + N$. As above, the generalized eigenvalues for

$$\begin{pmatrix} F & B^t \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \lambda \begin{pmatrix} F & 0 \\ 0 & \frac{1}{\nu} Q \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} \tag{2.6}$$

are either $\lambda = 1$ or (2.4), where μ is now a solution to the generalized eigenvalue problem

$$Sp = \mu \left(\frac{1}{\nu} Q \right) p, \tag{2.7}$$

with $S = BF^{-1}B^t$, the Schur complement for the discrete Oseen operator. The following result, which generalizes the analysis for the Stokes operator in [19], provides a bound.

THEOREM 1. *The eigenvalues of the generalized Schur complement problem* (2.7) *for the Oseen operator are contained in a rectangular box in the right half plane whose borders are bounded independently of h.*

Proof. Let $C = B\left(\frac{F^{-1}+F^{-t}}{2}\right)B^t$ denote the symmetric part of S, and let $R = B\left(\frac{F^{-1}-F^{-t}}{2}\right)B^t$ denote its skew-symmetric part, so that $S = C + R$. By Bendixson's theorem ([17, p. 418]), any eigenvalue μ of the problem (2.7) satisfies

$$\min_p \frac{(p, Cp)}{(p, \frac{1}{\nu} Qp)} \le \operatorname{Re}(\mu) \le \max_p \frac{(p, Cp)}{(p, \frac{1}{\nu} Qp)}, \qquad |\operatorname{Im}(\mu)| \le \max_p \frac{|(p, Rp)|}{(p, \frac{1}{\nu} Qp)}. \tag{2.8}$$

To construct bounds on these Rayleigh quotients, it will be convenient to refer to $S_\infty = BA^{-1}B^t$, the Schur complement for the Stokes operator. For the symmetric part C in (2.8), we use the relation

$$\frac{(p, Cp)}{(p, \frac{1}{\nu} Qp)} = \frac{(p, Cp)}{(p, \frac{1}{\nu} S_\infty p)} \frac{(p, S_\infty p)}{(p, Qp)}. \tag{2.9}$$

In light of (1.4), we need only consider the first quotient on the right in (2.9). Note that

$$\begin{aligned} \frac{F^{-1} + F^{-t}}{2} &= F^{-1}\left(\frac{F + F^t}{2}\right)F^{-t} \\ &= (\nu A + N)^{-1} (\nu A) (\nu A - N)^{-1} \\ &= A^{-1/2} \left(\nu I - \frac{1}{\nu} \tilde{N}^2\right)^{-1} A^{-1/2}, \end{aligned} \tag{2.10}$$

where $\tilde{N} = A^{-1/2} N A^{-1/2}$. Consequently,

$$\frac{(p, Cp)}{(p, \frac{1}{\nu} S_\infty p)} = \frac{(p, BA^{-1/2} (\nu I - \frac{1}{\nu} \tilde{N}^2)^{-1} A^{-1/2} B^t p)}{(p, \frac{1}{\nu} BA^{-1} B^t p)} = \frac{(v, (I - \frac{1}{\nu^2} \tilde{N}^2)^{-1} v)}{(v, v)},$$

where $v = A^{-1/2} B^t p$. But $\tilde{N}$ is skew-symmetric, so that the eigenvalues of $-\tilde{N}^2$ are real and nonnegative. Moreover, since N and A are first-order and second-order operators, respectively, the eigenvalues of $\tilde{N}$ are uniformly bounded in modulus by a constant δ that is independent of h [5]. Therefore, the spectrum of $I - \frac{1}{\nu^2} \tilde{N}^2$ is contained in the interval $\left[1, 1 + \delta^2/\nu^2\right]$, or, equivalently,

$$\frac{\nu^2}{\delta^2 + \nu^2} \le \frac{(p, Cp)}{(p, \frac{1}{\nu} S_\infty p)} \le 1.$$

Combining this with (1.4) and (2.9) gives

$$\frac{\gamma^2 \nu^2}{\delta^2 + \nu^2} \le \frac{(p, Cp)}{(p, \frac{1}{\nu} Qp)} \le \Gamma^2.$$

For the skew-symmetric part R in (2.8), the analogue of (2.9) is

$$\frac{(p, Rp)}{(p, \frac{1}{\nu}Qp)} = \frac{(p, Rp)}{(p, \frac{1}{\nu}S_\infty p)} \frac{(p, S_\infty p)}{(p, Qp)},$$

and as in (2.10), we have

$$\frac{F^{-1} - F^{-t}}{2} = -A^{-1/2}(\nu I + \tilde{N})^{-1}\tilde{N}(\nu I - \tilde{N})^{-1}A^{-1/2}.$$

Therefore

$$\frac{(p, Rp)}{(p, \frac{1}{\nu}S_\infty p)} = -\frac{\nu(v, \tilde{N}v)}{(v, (\nu^2 I - \tilde{N}^2)v)}, \tag{2.11}$$

where $v = (\nu I - \tilde{N})^{-1}A^{-1/2}B^T p$. The skew-symmetric matrix $\tilde{N}$ admits a decomposition of the form $\tilde{N} = iU\Lambda U^H$, where Λ is a real diagonal matrix and U is unitary. Consequently, $\tilde{N}^2 = -U\Lambda^2 U^H$, and the modulus of the Rayleigh quotient on the right side of (2.11) can be expressed in the form

$$\frac{\nu\,|(w, \Lambda w)|}{(w, (\nu^2 I + \Lambda^2)w)}.$$

This is bounded by

$$\max_{-\delta\le\lambda\le\delta} \frac{\nu|\lambda|}{\nu^2 + \lambda^2} = \max_{0\le\lambda\le\delta} \frac{\nu\lambda}{\nu^2 + \lambda^2}.$$

It follows from elementary calculus that this maximum is $\frac{1}{2}$, obtained when $\lambda = \nu$, giving

$$\frac{|(p, Rp)|}{(p, \frac{1}{\nu}Qp)} \le \frac{\Gamma^2}{2}. \qquad \square$$

Corollary 1 follows immediately from Theorem 1 and (2.4).

COROLLARY 1. *The eigenvalues of the discrete Oseen operator preconditioned by* (2.5) *consist of* $\lambda = 1$ *of multiplicity* $n_u - n_p$*, together with four sets consisting of points of the form* $1 + (a \pm bi)$ *and* $-a \pm bi$*. These sets can be enclosed in two rectangular regions that are symmetric with respect to* $\mathrm{Re}(\lambda) = \frac{1}{2}$*, whose borders are bounded independently of* h*.*

The inclusion regions for these eigenvalues consist of the image of the box $\left[\frac{\gamma^2\nu^2}{\delta^2+\nu^2}, \Gamma^2\right] \times \left[-\frac{\Gamma^2}{2}, \frac{\Gamma^2}{2}\right]$ under the mapping $\mu \mapsto \lambda(\mu)$ given by (2.4). It can be shown that the rectangular regions of this result are contained in

$$\left[\frac{1 + s_{\min}}{2}, \frac{1 + s_{\max}}{2}\right] \times [-t, t] \quad \text{and} \quad \left[\frac{1 - s_{\max}}{2}, \frac{1 - s_{\min}}{2}\right] \times [-t, t]$$

in the right and left half sides, respectively, of the complex plane, where

$$s_{\min} = \left(1 + \frac{4\gamma^2\nu^2}{\delta^2 + \nu^2}\right)^{1/2}, \qquad s_{\max} = \left[\frac{1}{2}\left(1 + 4\Gamma^2 + \sqrt{1 + 8\gamma^2 + 20}\,\right)\right]^{1/2},$$

$$t = \frac{\Gamma^2}{\left(1 + \frac{4\gamma^2\nu^2}{\delta^2+\nu^2}\right)^{1/2}}.$$

The fact that the eigenvalues for the preconditioned system derived from (2.5) lie on both sides of the imaginary axis is a potential disadvantage of this idea. An alternative that avoids this problem is the block triangular operator

$$\begin{pmatrix} F & B^t \\ 0 & -\frac{1}{\nu}Q \end{pmatrix}. \tag{2.12}$$

For this choice, the preconditioned eigenvalue problem is

$$\begin{pmatrix} F & B^t \\ B & O \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix} = \lambda \begin{pmatrix} F & B^t \\ 0 & -\frac{1}{\nu}Q \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix}. \tag{2.13}$$

Again, one solution is $\lambda = 1$, now of multiplicity n_u. If $\lambda \neq 1$, then premultiplying the first block row of (2.13) by BF^{-1} and using the relation $Bu = -\lambda(\frac{1}{\nu}Q)p$ leads to equation (2.7) for the other eigenvalues. Thus, we have the following result.

THEOREM 2. *The eigenvalues of the discrete Oseen operator preconditioned by* (2.12) *consist of* $\lambda = 1$ *together with the generalized eigenvalues of* S *in* (2.7). *Therefore, the eigenvalues are bounded independently of* h.

Remark 1. Use of either preconditioning operator (2.5) or (2.12) entails the computation of the action of F^{-1} at each step of an iterative procedure. F is a discrete convection-diffusion operator and applying F^{-1} to a vector using direct methods will be expensive. An alternative is to replace this computation with an approximation obtained by iterative solution of the convection-diffusion equation. We will examine this approach in §4.

Remark 2. Both preconditioners also require the action of Q^{-1}, which may also be expensive, depending on the choice of pressure discretization. In this case, however, it is known that Q can be replaced by some approximation $\hat{Q}$ without affecting asymptotic convergence properties; only the constants γ and Γ of (1.4) change [21]. In the experiments discussed in §§3 and 4, we replace Q with a diagonal matrix consisting of the main diagonal of Q.

3. Numerical results I: Exact convection-diffusion solves. In this section, we present the results of numerical experiments indicating that the analysis of §2 is predictive of the performance of iterative methods for solving (1.3). Computations were performed using MATLAB 4.1 on a variety of computing platforms.

Our test problem is a "leaky" two-dimensional lid-driven cavity problem in a square domain $(-1 \leq x \leq 1 : -1 \leq y \leq 1)$. The boundary conditions are $u_x = u_y = 0$ on the three fixed walls $(x = -1, y = -1, x = 1)$, and $u_x = 1$, $u_y = 0$ on the moving wall $(y = 1)$. The hydrostatic pressure is not explicitly specified, so that all the linear equation systems we solve below are singular with a one-dimensional nullspace. The convective "wind" is a circular vortex as illustrated in Fig. 1, and is given by

$$\begin{aligned} w_x &= 2y(1 - x^2), \\ w_y &= -2x(1 - y^2). \end{aligned}$$

The fact that there is no dominant flow direction makes this a challenging test problem. Note that in the corners and in the center of the flow region the driving flow is stagnant.[1]

Unless otherwise specified, we consider values of the viscosity parameter ν between 1 and 1/100. When $\nu = 1$ we have diffusion dominated (essentially Stokes) flow, whereas as

[1] This is actually not a linearization of the Navier–Stokes equations since **w** and **u** do not satisfy the same boundary conditions; this fact does not affect the demonstration of the solution algorithms. Also, note that div **w** = 0 so that the term $\frac{1}{2}\mathbf{u}(\operatorname{div}\mathbf{w})$ in (1.2) is identically zero.

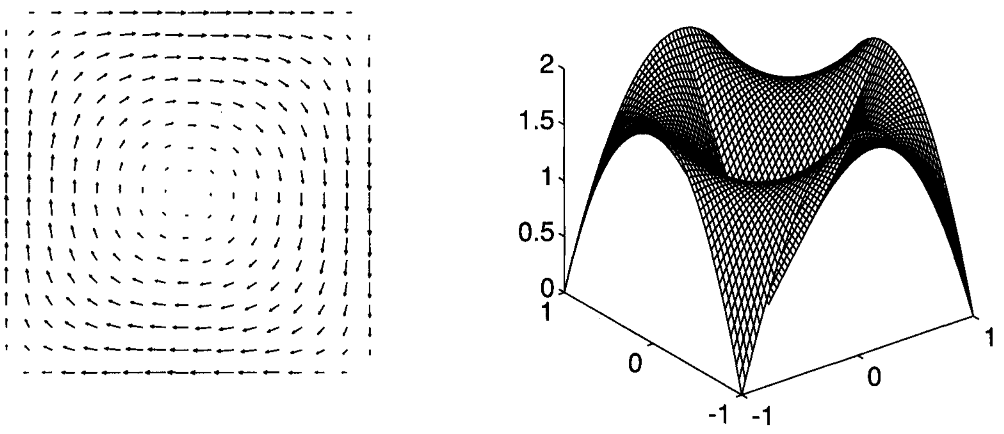

FIG. 1. *Magnitude and direction of the convecting flow.*

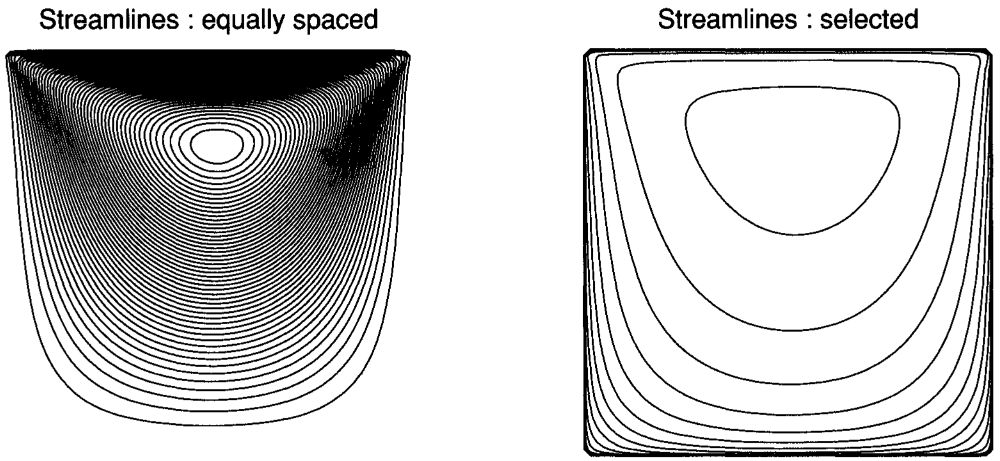

FIG. 2. *Uniform* 64 × 64 *grid*: $\nu = 1$.

$\nu \to 0$ the flow becomes dominated by the "wind." Typical flow solutions are illustrated in Figs. 2 and 3. Note that as the viscosity is decreased, the center of primary recirculation moves to the right (the Stokes flow solution is perfectly symmetric about the line $x = 0$), and secondary vortices are generated in the two bottom corners.

To discretize (1.2), we take a finite element subdivision based on $n \times n$ grids of rectangular elements. Bearing in mind the nature of the flow solution being computed, we present results for two representative discretizations here: a conventional Galerkin approach using a quasi-uniform sequence of grids and a streamline-upwind method using uniform grids of square elements of size $h = 2/n$. In either case, the mixed finite element used was the div-stable "Taylor–Hood" method based on continuous bilinear pressure with a continuous bilinear velocity field defined on four element macroelements (see, e.g., [10, p. 30]).

For the Galerkin discretization, the quasi-uniform grids are chosen to resolve the details of the flow in the four corners of the domain: they are symmetric about $x = 0$ and $y = 0$, and in each quadrant the grid lines become more dense near the boundary. The 64 × 64 grid is shown in the pressure solution plot in Fig. 4. The analytic pressure solution is singular at the top corners where the imposed velocity is discontinuous.

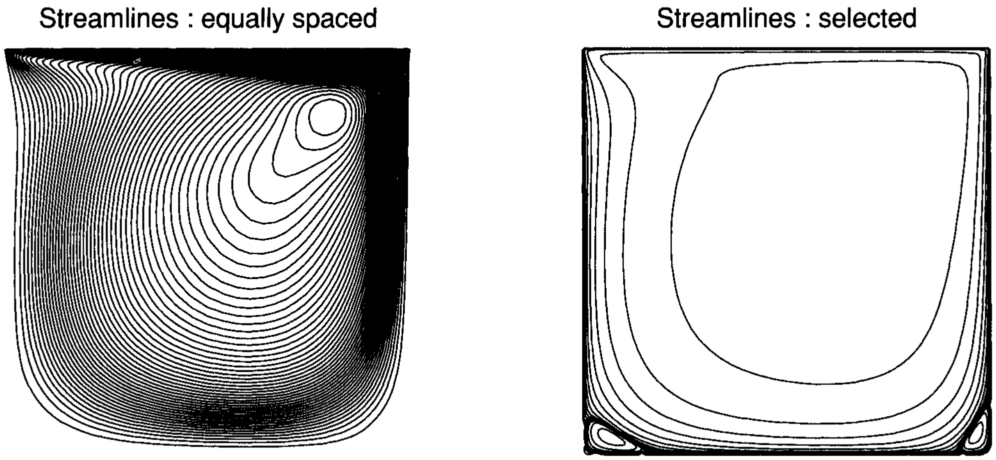

FIG. 3. *Nonuniform* 64 × 64 *grid*: $\nu = 1/100$.

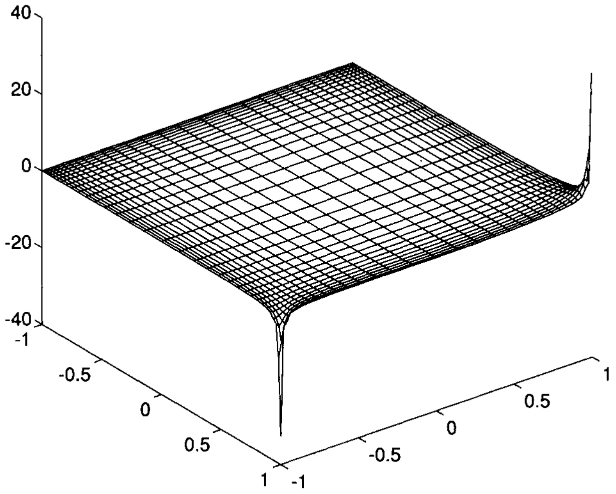

FIG. 4. *Pressure solution for* $\nu = 1/10$.

The streamline-upwind discretization is as described in [12, p. 185]. In this case, the block convection-diffusion operator F is perturbed by a symmetric positive semidefinite matrix A_{w}. That is, $F = -\nu\Delta_h + A_{\mathrm{w}} + N$, where Δ_h is the discrete Laplacian obtained from the usual Galerkin formulation. A_{w} is the discrete form of a stabilizing term $\alpha\,(\mathbf{w}\cdot\nabla\mathbf{u}, \mathbf{w}\cdot\nabla\mathbf{v})$ that adds $\alpha \equiv O(h)$ diffusion along the streamlines. For our experiments with streamline upwinding, we took $\alpha = h/4$. Note that the perturbation does not affect the skew-symmetric part of the convection-diffusion operator, so that the analysis of §2 holds; only the definition of the "diffusion matrix" A is changed, from $-\Delta_h$ to $-\Delta_h + \frac{1}{\nu}A_{\mathrm{w}}$.

We first consider the bounds of Theorem 1. Table 1 shows the extreme real parts and maximum imaginary parts of the generalized eigenvalues (2.7) of the Schur complement operator, for $\nu = 1/10$ and $1/100$ with the streamline-upwind discretization, on three meshes. The small changes in all values are in accordance with the analysis, although it appears that finer meshes would be needed to produce constant values. The analysis also shows that the

TABLE 1
Eigenvalues of the Schur complement for streamline-upwind discretization.

	$\nu = 1/10$			$\nu = 1/100$		
Grid	Min Re	Max Re	Max Im	Min Re	Max Re	Max Im
16 × 16	7.17E-2	1.11	0.46	1.66E-2	1.07	0.20
32 × 32	8.75E-2	1.64	0.71	1.33E-2	1.11	0.50
64 × 64	9.08E-2	2.00	0.87	1.14E-2	1.37	0.74

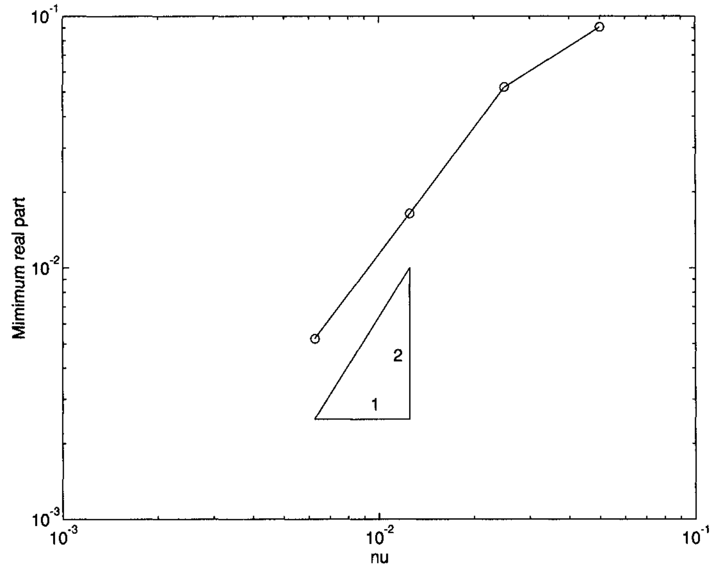

FIG. 5. *Minimum real parts of eigenvalues of the Schur complement, for streamline-upwind discretization on a* 64 × 64 *grid.*

real parts and largest imaginary parts of the eigenvalues are bounded independently of ν; the bound for the smallest real part is proportional to ν^2. The data in Table 1 are in agreement with the upper bounds. Figure 5 plots the smallest real parts on a logarithmic scale for the streamline-upwind discretization on a 64 × 64 grid and $\nu = 1/20$, $1/40$, $1/80$, and $1/160$. The results indicate that the lower bound is also tight.

We test the preconditioners here with two Krylov subspace methods for solving nonsymmetric systems: the generalized minimum residual method (GMRES) [15] and a simple implementation of the quasi-minimum residual method (QMR) [8] based on coupled two-term recurrences without look-ahead. GMRES demonstrates the performance of the preconditioners with the optimal (with respect to the residual norm) Krylov subspace solver. This method is impractical for large problems because its work and storage requirements grow with the iteration count; QMR is a nonoptimal alternative that avoids this difficulty. Some additional experiments with restarted GMRES are presented in §4. In all cases we use right-oriented preconditioning, and our convergence criterion is a reduction of 10^{-6} in the l_2-norm of the residual. The action of F^{-1} and F^{-t} is computed using the LU-factorization in MATLAB. We start from a zero initial guess. Using random initial guesses gave comparable iteration counts in all cases.

We first discuss the performance of GMRES. Table 2 shows the iteration counts obtained for three values of ν using the block triangular preconditioner (2.12) in the case of Galerkin approximation on the quasi-uniform grid sequence, and Table 3 shows the iteration counts for uni-

TABLE 2
GMRES iteration counts for Galerkin discretization with block triangular preconditioner.

Grid	16 x 16	32 x 32	64 x 64	128 x 128
$\nu = 1$	18	19	17	14
$\nu = 1/10$	25	31	32	31
$\nu = 1/50$	45	69	93	110

TABLE 3
GMRES iteration counts for streamline-upwind discretization with block triangular preconditioner.

Grid	16 x 16	32 x 32	64 x 64	128 x 128
$\nu = 1$	21	22	21	19
$\nu = 1/10$	32	36	35	33
$\nu = 1/50$	48	72	97	111

TABLE 4
QMR iteration counts for Galerkin discretization with block triangular (diagonal) preconditioner.

Grid	16 x 16	32 x 32	64 x 64	128 x 128
$\nu = 1$	22 (43)	22 (43)	22 (41)	16
$\nu = 1/10$	28 (51)	36 (68)	39 (78)	36
$\nu = 1/50$	51 (100)	78 (155)	112 (223)	127
$\nu = 1/100$	73 (143)	126 (246)	189 (375)	253

form grids with streamline-upwind discretization. The results suggest that grid-independent convergence is observed even on relatively coarse grids if the flow is diffusion dominated, that is, eigenvalues are indeed predictive of performance. If convection dominates (as ν tends to zero), then the iteration counts increase, as might be expected from the analytic bounds of §2. For the smallest value of ν considered here, 1/50, the iteration counts grow as the mesh is refined; although for fine enough grids the counts become close to constant. We believe that for "small" ν, the asymptotic grid independence will be visible only for sufficiently fine grids.

GMRES with the block diagonal preconditioner (2.5) gives an identical picture. Indeed, we find that for the grid sizes and values of ν in Tables 2 and 3 (only $n \leq 64$ was tested), GMRES requires precisely $2k - 1$ iterations to reach the tolerance where k is the iteration count from Table 2 or 3. Moreover, the behavior of GMRES using (2.5) mirrors its behavior with the triangular preconditioning, in the sense that the odd iterates at step $2i - 1$ are close to those obtained in the triangular case at step i, and the norms of the even iterates stagnate. A rigorous explanation for this behavior can be given by relating the optimal polynomials implicitly generated by GMRES. In particular, it follows from the analysis in [6] that for a particular starting guess (not zero), the ith GMRES polynomial for the triangular case is identical to the $(2i - 1)$st GMRES polynomial for the diagonal case.

Tables 4 and 5 show analogous iteration counts for QMR. The tables contain data for both the triangular and (in parentheses for $n \leq 64$) diagonal preconditioners. Note that the iteration counts in the first three rows of Tables 4 and 5 are close to the corresponding entries of Tables 2 and 3, respectively; i.e., the performance in QMR is close to optimal. Results for $\nu = 1/100$ are included to get a sense of the behavior of the preconditioners as ν becomes small; it appears that the asymptotic behavior is not seen for the grid sizes used.

Note that the cost per step of the block triangular preconditioner is only slightly higher than that of the block diagonal preconditioner (only an extra multiplication by B^t is needed); hence (2.12) is more efficient. However, for the Stokes problem ($\nu \to \infty$), the preferred choice of preconditioner is less obvious since the inherent symmetry is destroyed if (2.12) is used in place of (2.5).

TABLE 5

QMR iteration counts for streamline-upwind discretization with block triangular (diagonal) preconditioner.

Grid	16 × 16	32 × 32	64 × 64	128 × 128
$\nu = 1$	25 (49)	27 (51)	25 (47)	22
$\nu = 1/10$	36 (78)	44 (91)	42 (80)	37
$\nu = 1/50$	61 (127)	91 (175)	117 (234)	130
$\nu = 1/100$	76 (157)	133 (249)	190 (382)	229

Remark 3. In addition to the implementation of QMR with a coupled two-term recurrence (QMR$_2$) discussed above, we tested a version without look-ahead based on a three-term recurrence (QMR$_3$) [7], and the definitive (Fortran) implementation of a two-term QMR with look-ahead (QMR$_2^*$) from the QMRPAK directory in Netlib. For the preconditioners discussed above, the performances of the three variants were virtually identical. However, with the inexact preconditioners of the next section, we found QMR$_2$ to be much more robust than QMR$_3$.

4. Numerical results II: Inexact convection-diffusion solves. The dominant costs of the preconditioners of §§2 and 3 come from applying the action of F^{-1}, and for QMR, F^{-t}, to some vector v at each step of the iteration. In this section, we show that this operation can be replaced by an inexpensive one derived from an approximation to F^{-1}, with little degradation of performance of the Krylov subspace methods. The idea is to replace the preconditioning operators (2.5) and (2.12) with

$$(4.1) \qquad \begin{pmatrix} \hat{F} & 0 \\ 0 & \frac{1}{\nu}Q \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \hat{F} & B^t \\ 0 & -\frac{1}{\nu}Q \end{pmatrix},$$

respectively, where $\hat{F} \approx F$. Our choice of $\hat{F}$ will be implicitly determined by the use of iterative methods to compute approximate solutions to the systems $Fw = v$ and $F^tw = v$, although the methodology is not restricted to this choice. We will refer to the preconditioners that use the exact action of F^{-1} as the *exact* versions, and those based on approximations as in (4.1) as *inexact* versions.

Some insight into the effects of the inexact preconditioners is derived from matrix perturbation theory. Let $Q_\nu = \frac{1}{\nu}Q$. The preconditioned matrix for the exact block diagonal preconditioner (2.5) is

$$\mathcal{A}_D = \begin{pmatrix} F & B^t \\ B & O \end{pmatrix} \begin{pmatrix} F & 0 \\ 0 & Q_\nu \end{pmatrix}^{-1} = \begin{pmatrix} I & B^tQ_\nu^{-1} \\ BF^{-1} & 0 \end{pmatrix},$$

and that derived from the inexact version is

$$\hat{\mathcal{A}}_D = \begin{pmatrix} F & B^t \\ B & O \end{pmatrix} \begin{pmatrix} \hat{F} & 0 \\ 0 & Q_\nu \end{pmatrix}^{-1} = \mathcal{A}_D + \mathcal{E}_D,$$

where, with $E = \hat{F} - F$,

$$\mathcal{E}_D = -\begin{pmatrix} E\hat{F}^{-1} & 0 \\ BF^{-1}E\hat{F}^{-1} & 0 \end{pmatrix}.$$

Similarly, the preconditioned matrices for the exact and inexact block tridiagonal preconditioners (2.12) satisfy $\hat{\mathcal{A}}_T = \mathcal{A}_T + \mathcal{E}_T$, where

$$\mathcal{A}_T = \begin{pmatrix} I & 0 \\ BF^{-1} & BF^{-1}B^tQ_\nu^{-1} \end{pmatrix}, \qquad \mathcal{E}_T = -\begin{pmatrix} E\hat{F}^{-1} & E\hat{F}^{-1}B^tQ_\nu^{-1} \\ BF^{-1}E\hat{F}^{-1} & BF^{-1}E\hat{F}^{-1}B^tQ_\nu^{-1} \end{pmatrix}.$$

We have the following bounds on the eigenvalues of the preconditioned systems using inexact preconditioners.

THEOREM 3. *If $\mathcal{A}_D = \mathcal{V}_D \Lambda_D \mathcal{V}_D^{-1}$ is diagonalizable, then for any eigenvalue $\mu \in \sigma(\hat{\mathcal{A}}_D)$,*

$$\min_{\lambda \in \sigma(\mathcal{A}_D)} |\lambda - \mu| \leq \|E\hat{F}^{-1}\|_\infty \kappa_\infty(\mathcal{V}_\mathcal{D}) \max(1, \|BF^{-1}\|_\infty).$$

If $\mathcal{A}_T = \mathcal{V}_T \Lambda_T \mathcal{V}_T^{-1}$ is diagonalizable, then for any eigenvalue $\mu \in \sigma(\hat{\mathcal{A}}_T)$,

$$\min_{\lambda \in \sigma(\mathcal{A}_T)} |\lambda - \mu| \leq \|E\hat{F}^{-1}\|_\infty \kappa_\infty(\mathcal{V}_T)\,(1 + \|B^t Q^{-1}\|_\infty) \max(1, \|BF^{-1}\|_\infty).$$

Proof. The result is an immediate consequence of the Bauer–Fike theorem [9, p. 342], which states that for diagonalizable $\mathcal{A} = \mathcal{V}\Lambda\mathcal{V}^{-1}$, any $\mu \in \sigma(\mathcal{A} + \mathcal{E})$ satisfies $\min_{\lambda \in \sigma(\mathcal{A})} |\lambda - \mu| \leq \kappa(\mathcal{V})\, \|\mathcal{E}\|$, where $\|\cdot\|$ is any l_p-norm. □

Thus, if $\hat{F}$ is a good enough approximation to F, i.e., if enough inner iterations are used, then $\|E\hat{F}^{-1}\|$ will be small and the eigenvalues of $\hat{\mathcal{A}}_D$ and $\hat{\mathcal{A}}_T$ will be close to those of $\mathcal{A}_D$ and $\mathcal{A}_T$, respectively. We state the result in terms of the l_∞-norm only because the bounds then have a simple form.

Remark 4. We have computed the condition numbers $\kappa(\mathcal{V})$ for $\mathcal{A}_T$ and found them to be large, on the order of 10^3 or higher, for the three values of ν, with streamline upwinding and $h = 1/16$. However, the presence of $\kappa(\mathcal{V})$ in these bounds is an artifact of the proof of the Bauer–Fike theorem; there are more subtle analyses ([9, pp. 344ff]), as well as bounds that do not require diagonalizable matrices [11]. We have observed that the eigenvalues of $\mathcal{A}_T$ are insensitive to perturbations, and we believe that the presence of $\kappa(\mathcal{V})$ is pessimistic. This supposition is supported by the experimental results described below.

To demonstrate that inexact preconditioning is effective, we consider versions of it based on two line-oriented splittings of F. The first uses a *horizontal line Gauss–Seidel relaxation*: Let $F = H - R$ denote a horizontal line Gauss–Seidel splitting of the block convection-diffusion operator F derived from the one-line natural left-to-right, bottom-to-top ordering of the velocity grid. Thus, H is a block lower triangular matrix consisting of the block diagonal of F (a tridiagonal matrix) together with the strict block lower triangular part of F. (See [18], [22] for further details.) The horizontal line Gauss–Seidel method for $Fw = v$ performs the iteration

$$w_0 = 0, \quad w_{i+1} = w_i + H^{-1}(v - Fw_i).$$

For k steps of this iteration, the approximating matrix is $\hat{F} = F(I - (H^{-1}R)^k)^{-1}$.

It has been observed that the performance of relaxation methods of this type can be improved if the sweep direction follows the underlying direction of flow [4]. Our benchmark problem has a circular flow, so that no simple line relaxation can mimic the flow direction throughout Ω. A slightly more sophisticated idea is to use an *alternating line relaxation.* For this, let $F = V - T$ denote a vertical line Gauss–Seidel splitting of F; that is, if P is a permutation matrix associated with the mapping from the natural horizontal line ordering of grid points to the natural vertical line ordering, then $P^T VP$ is the block lower triangular part of $P^T FP$. One iteration of alternating line relaxation consists of two line Gauss–Seidel steps, one using the horizontal splitting, followed by one using the vertical splitting:

$$w_0 = 0, \quad w_{i+1/2} = w_i + H^{-1}(v - Fw_i), \quad w_{i+1} = w_{i+1/2} + V^{-1}(v - Fw_{i+1/2}).$$

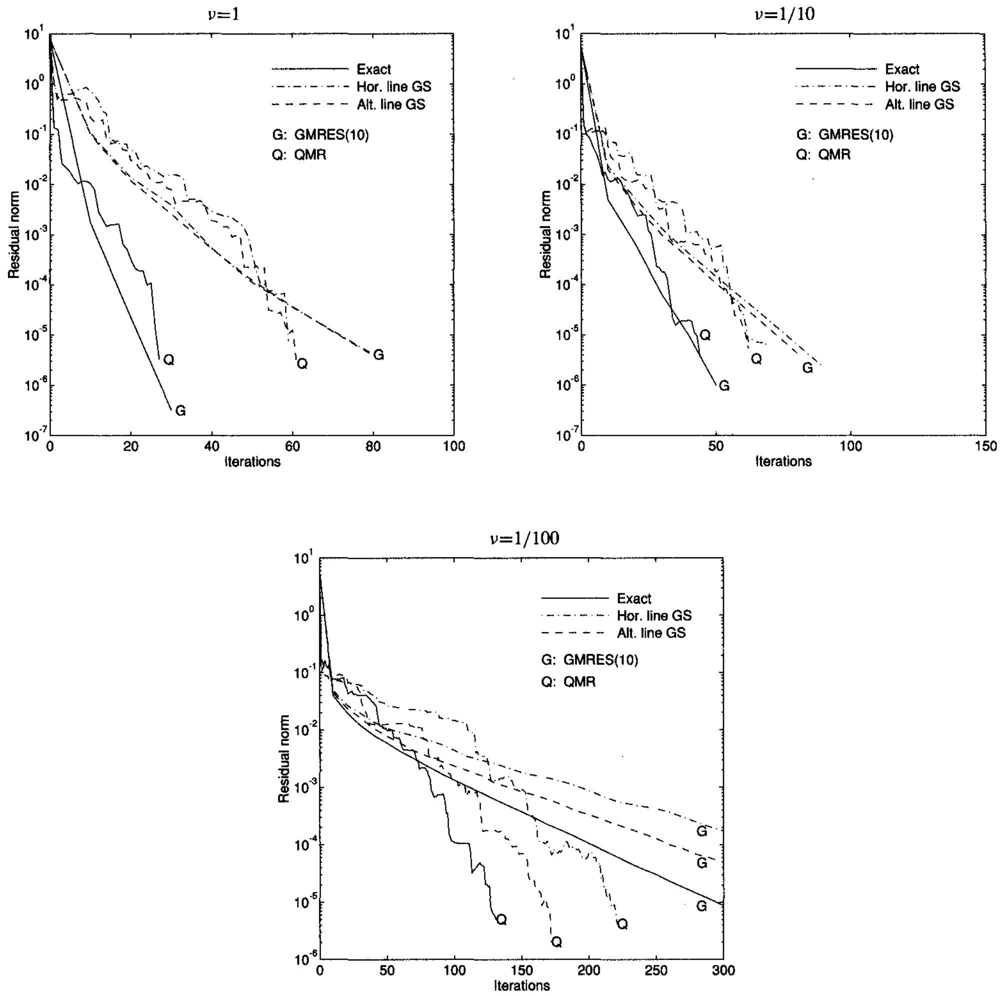

FIG. 6. *Performance of block tridiagonal inexact preconditioners for* 32 × 32 *grid.*

Figure 6 shows results of using the inexact block tridiagonal preconditioners with a version of GMRES with restarts every ten steps (denoted GMRES(10)) and with QMR. The test problem is discretized by streamline upwinding on a 32 × 32 grid. Results for inexact block diagonal preconditioners were similar, except that, as with the exact preconditioners, convergence was slower. We used four steps of horizontal line relaxation or two steps of alternating line relaxation, so that both inexact preconditioners perform four sweeps. The figure also shows the performance of the exact preconditioner, whose cost per step is significantly more expensive. For example, with an $n \times n$ velocity grid, direct solution using a bandsolver requires $O(n^4)$ operations, whereas each inner iteration is an $O(n^2)$ computation. We see that the use of inexact preconditioners in place of the exact versions leads to little degradation of performance of the Krylov subspace methods. For example, in the convection-dominated case $\nu = 1/100$, QMR with alternating line relaxation requires roughly 25% more iterations than with the exact preconditioner. For the diffusion dominated case $\nu = 1$, roughly three times as many outer iterations are required with the inexact preconditioners, which still leads to a less costly computation. Not surprisingly, alternating relaxation is more effective than horizontal relaxation, especially for convection-dominated problems. We remark that our goal here is only to demonstrate "proof-of-concept"; many other techniques for approximating the action of F^{-1} are possible, for both diffusion-dominated and convection-dominated flow. See, for example, [1], [4], [14], [16], [20].

Remark 5. Although we do not make a detailed comparison of Krylov subspace methods, we briefly comment on the behavior of the two choices used here. QMR requires twice as many preconditioned matrix–vector products per step as GMRES(10), and since matrix–vector products are the dominant cost, each QMR step is roughly twice as expensive. Thus, these results indicate that GMRES(10) is more efficient than QMR for large ν, but QMR becomes more effective as convection becomes dominant. The storage requirements of the two methods are comparable.

Acknowledgments. We thank Oliver Ernst, Roland Freund, and Marlis Hochbruck for providing us with Matlab versions of QMR.

REFERENCES

[1] O. AXELSSON, V. EIJKHOUT, B. POLMAN, AND P. VASSILEVSKI, *Incomplete block-matrix factorization iterative methods for convection-diffusion problems*, BIT, 29 (1989), pp. 867–889.

[2] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[3] A. BROOKS AND T. HUGHES, *Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations*, Comp. Methods Appl. Mech. Engrg., 32 (1982), pp. 199–259.

[4] H. C. ELMAN AND G. H. GOLUB, *Line iterative methods for cyclically reduced convection-diffusion problems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 339–363.

[5] H. C. ELMAN AND M. H. SCHULTZ, *Preconditioning by fast direct methods for nonselfadjoint nonseparable elliptic problems*, SIAM J. Numer. Anal, 23 (1986), pp. 44–57.

[6] B. FISCHER, A. RAMAGE, D. SILVESTER, AND A. J. WATHEN, *Minimum Residual Methods for Augmented Systems*, Tech. Report 15, Department of Mathematics, Strathclyde University, Glasgow, Scotland, 1995.

[7] R. FREUND AND N. M. NACHTIGAL, *QMR: A quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[8] ———, *An implementation of the QMR method based on coupled two-term recurrences*, SIAM J. Sci. Comput., 15 (1994), pp. 313–337.

[9] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, Baltimore, MD, 1989.

[10] M. GUNZBURGER, *Finite Element Methods for Viscous Incompressible Flows*, Academic Press, San Diego, 1989.

[11] P. HENRICI, *Bounds for iterates, inverses, spectral variation, and fields of values of non-normal matrices*, Numer. Math., 4 (1962), pp. 24–40.

[12] C. JOHNSON, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, New York, 1987.

[13] K. W. MORTON, A. PRIESTLEY, AND E. SÜLI, *Stability of the Lagrange-Galerkin method with non-exact integration*, Math. Modelling Numer. Anal., 22 (1988), pp. 625–653.

[14] T. RUSTEN AND R. WINTHER, *A preconditioned iterative method for saddle point problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.

[15] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[16] D. SILVESTER AND A. WATHEN, *Fast iterative solution of stabilised Stokes systems part II: Using block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367.

[17] J. STOER AND R. BULIRSCH, *Introduction to Numerical Analysis*, 2nd ed., Springer-Verlag, New York, 1993.

[18] R. S. VARGA, *Matrix Iterative Analysis*, Prentice–Hall, Englewood Cliffs, NJ, 1962.

[19] R. VERFÜRTH, *A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem*, IMA J. Numer. Anal., 4 (1984), pp. 441–455.

[20] A. WATHEN AND D. SILVESTER, *Fast iterative solution of stabilized Stokes systems. Part I: Using simple diagonal preconditioners*, SIAM J. Numer. Anal., 30 (1993), pp. 630–649.

[21] A. J. WATHEN, *Realistic eigenvalue bounds for the Galerkin mass matrix*, IMA J. Numer. Anal., 7 (1987), pp. 449–457.

[22] D. M. YOUNG, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1970.