Chapter 11
A Preconditioned Scheme for Nonsymmetric Saddle-Point Problems

Abdelkader Baggag

Abstract In this paper, we present an effective preconditioning technique for solving nonsymmetric saddle-point problems. In particular, we consider those saddle-point problems that arise in the numerical simulation of particulate flows—flow of solid particles in incompressible fluids, using mixed finite element discretization of the Navier–Stokes equations.

These indefinite linear systems are solved using a preconditioned Krylov subspace method with an indefinite preconditioner. This creates an inner–outer iteration, in which the inner iteration is handled via a preconditioned Richardson scheme. We provide an analysis of our approach that relates the convergence properties of the inner to the outer iterations. Also "optimal" approaches are proposed for the implicit construction of the Richardson's iteration preconditioner. The analysis is validated by numerical experiments that demonstrate the robustness of our scheme, its lack of sensitivity to changes in the fluid–particle system, and its "scalability".

11.1 Introduction

Many scientific applications require the solution of saddle-point problems of the form

$$\begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \tag{11.1}$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ with $m \leq n$, and where the $(n + m) \times (n + m)$ coefficient matrix

$$\mathscr{A} = \begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix},$$

is assumed to be nonsingular. Such systems are typically obtained when "Lagrange multipliers" or mixed finite element discretization techniques are employed. Examples of these include, but are not limited to, the equality-constrained quadratic pro-

A. Baggag (✉)
College of Science and Engineering, Université Laval, Quebec City, Canada
e-mail: abdelkader.baggag@gci.ulaval.ca

M.W. Berry et al. (eds.), *High-Performance Scientific Computing*,
DOI 10.1007/978-1-4471-2437-5_11,

gramming problems, the discrete equations which result from the approximation of elasticity problems, Stokes equations, and the linearization of Navier–Stokes equations [2, 15, 16, 29, 39, 47]. When the matrix A is symmetric and positive definite, the problem (11.1) has n positive and m negative eigenvalues, with well defined bounds [57]. If the matrix A is symmetric indefinite or nonsymmetric, little can be said about the spectrum of the indefinite matrix $\mathscr{A}$.

Much attention has been paid to the case when A is symmetric positive definite, e.g. see [1, 6, 8, 9, 11–13, 18, 22, 25–28, 32, 40, 46, 54–56, 64, 71–75], and more recently to the case when A is nonsymmetric [3, 4, 10, 14, 17, 19–21, 23, 24, 31, 41, 44, 45, 61–63, 65]. In this paper, A is assumed to be nonsymmetric and B of full column rank. Here, we adopt one of the symmetric indefinite preconditioners studied, among others, by Golub and Wathen [31], for solving Eq. (11.1) via a preconditioned Krylov subspace method, such as GMRES, with the preconditioner given by

$$\mathscr{M} = \frac{1}{2}\left(\mathscr{A} + \mathscr{A}^T\right) = \begin{bmatrix} A_s & B \\ B^T & 0 \end{bmatrix}. \tag{11.2}$$

Here, A_s is the symmetric part of A, i.e., $A_s = (A + A^T)/2$. The motivating application in our paper produces a block diagonal matrix A_s in which each block has the following properties:

1. positive definite and irreducibly diagonally dominant, i.e., for each diagonal block $A_s^{(k)} = [a_{ij}^{(s)}]$ is irreducible, and $a_{ii}^{(s)} \geq \sum_{j \neq i} |a_{ij}^{(s)}|$ with strict inequality holding for at least one i, and
2. $\| A_s \|_F \geq \| A_{ss} \|_F$ where $\| \cdot \|_F$ denotes the Frobenius norm, and $A_{ss} = (A - A^T)/2$ is the skew symmetric part of A.

Thus, the preconditioner $\mathscr{M}$ is nonsingular, and the Schur complement, $-(B^T A_s^{-1} B)$, is symmetric negative definite.

The application of the preconditioner $\mathscr{M}$ in each Krylov iteration requires the solution of a linear system of the form

$$\begin{bmatrix} A_s & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}. \tag{11.3}$$

The focus of our study is the development of a preconditioned Richardson iterative scheme for solving the above symmetric indefinite system (11.3) in a nested iterations setting that ensures the convergence of the inner iterations.

This system can be reformulated as

$$A_s \mathbf{x} = \mathbf{f} - B\mathbf{y}, \tag{11.4}$$

$$\left(B^T A_s^{-1} B\right)\mathbf{y} = B^T A_s^{-1}\mathbf{f} - \mathbf{g}. \tag{11.5}$$

Thus, one may first solve Eq. (11.5) to obtain $\mathbf{y}$, then solve Eq. (11.4) to get $\mathbf{x}$. Using a conjugate gradient algorithm for solving Eqs. (11.5), and (11.4), one creates an inner–outer iterative scheme [11]. This is the approach used in the classical

Uzawa scheme [1]. It turns out that in order to ensure convergence of the outer iteration, it is necessary to solve systems in the inner iteration with relatively high accuracy [13, 22]. For large-scale applications, such as the numerical simulation of particulate flows, solving linear systems involving A_s or $(B^T A_s^{-1} B)$ is not practical, as the action of A_s^{-1} must be computed on various vectors. Consequently, the approach we adopt here is to replace the cost of computing the action of A_s^{-1} by the cost of evaluating the action of some other "more economical" symmetric positive definite operator $\hat{A}^{-1}$ which approximates A_s^{-1} in some sense. Thus, the linear system (11.4) is solved via the iteration

$$\mathbf{x}_{k+1} = \left(I - \hat{A}^{-1} A_s\right) \mathbf{x}_k + \hat{A}^{-1} \underline{\mathbf{f}}, \tag{11.6}$$

where $\underline{\mathbf{f}} = \mathbf{f} - B\,\mathbf{y}$ and $\hat{A}$ is an appropriate symmetric positive definite splitting that assures convergence, i.e., $\alpha = \rho(I - \hat{A}^{-1} A_s) < 1$, where $\rho(\cdot)$ is the spectral radius.

Similarly, we replace A_s by $\hat{A}$ in (11.5) and solve the resulting "inexact" system,

$$\left(B^T \hat{A}^{-1} B\right) \mathbf{y} = B^T \hat{A}^{-1} \mathbf{f} - \mathbf{g}, \tag{11.7}$$

instead of the original system Eq. (11.5), via the iteration

$$\mathbf{y}_{k+1} = \left[I - \hat{G}^{-1}\left(B^T \hat{A}^{-1} B\right)\right] \mathbf{y}_k + \hat{G}^{-1} \hat{\mathbf{s}}, \tag{11.8}$$

where $\hat{\mathbf{s}} = B^T \hat{A}^{-1} \mathbf{f} - \mathbf{g}$, and $\hat{G}^{-1}$ is an inexpensive symmetric positive definite approximation of the inverse of the inexact Schur complement $\left(B^T \hat{A}^{-1} B\right)^{-1}$ that assures convergence of Eq. (11.8), i.e., $\beta = \rho(I - \hat{G}^{-1}(B^T \hat{A}^{-1} B)) < 1$. Moreover, $\hat{G}^{-1}$ is chosen such that $(I - \hat{G}^{-\frac{1}{2}}(B^T \hat{A}^{-1} B)\, \hat{G}^{-\frac{1}{2}})$ is positive definite.

Similarly, if we define the symmetric preconditioner $\hat{\mathscr{M}}$ to the system (11.3) as

$$\hat{\mathscr{M}} = \begin{bmatrix} \hat{A} & B \\ B^T & -\hat{G} + (B^T \hat{A}^{-1} B) \end{bmatrix}, \tag{11.9}$$

we obtain the following preconditioned Richardson iterative scheme for solving Eq. (11.3):

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} + \begin{bmatrix} \hat{A} & B \\ B^T & -\hat{G} + B^T \hat{A}^{-1} B \end{bmatrix}^{-1} \left\{ \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix} - \begin{bmatrix} A_s & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} \right\}, \tag{11.10}$$

that is convergent if and only if $\rho(\mathscr{I} - \hat{\mathscr{M}}^{-1} \mathscr{M}) < 1$.

Thus, our proposed nested iterative scheme is shown in Fig. 11.1 in which the outermost iteration is that of a Krylov subspace method (we use *restarted* GMRES throughout this paper), and the preconditioning operation itself is doubly nested. Our focus here is the development of an algorithm for the most inner iteration, i.e. solving systems involving the symmetric indefinite preconditioner (11.2) using the preconditioned Richardson iteration (11.10).

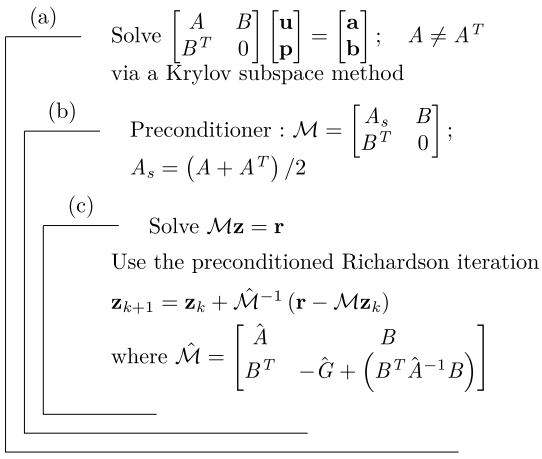


Fig. 11.1 A nested iterative scheme

In the Golub–Wathen study [31], an iteration of the form

$$\begin{bmatrix} A_s & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_{k+1} \\ \mathbf{p}_{k+1} \end{bmatrix} = \begin{bmatrix} (A_s - A) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{p}_k \end{bmatrix} + \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix},$$

is used, which does not always converge, and when used as an inner iteration within *full* GMRES, systems of the form $\mathscr{M}\mathbf{z} = \mathbf{r}$, in the inner-most loop $\{c\}$ of Fig. 11.1, are solved using a direct scheme. This could be as time-consuming as solving directly the nonsymmetric system (11.1), especially for very large systems.

In our study, the *monotone* convergence of our inner iteration (11.10) is guaranteed, and the performance of our nested scheme in Fig. 11.1 does not degrade as the mesh size decreases. Moreover, the construction of the preconditioner $\hat{\mathscr{M}}$ of the Richardson iteration is simple and economical.

In this paper, we analyze the iterative scheme (11.10) and show that a sufficient condition for *monotone* convergence is $\max\{\alpha, \beta\} < (\sqrt{5} - 1)/2$, and thus relating the rate of convergence of the inner iterations to the outer iteration, even though Eq. (11.8) is not the iteration that corresponds to the exact system (11.5) to be solved but to a modified one, (11.7), which, we will show, is not required to be solved accurately.

We use a simple explicit approximate inverse A_0^{-1} of A_s^{-1} for which $\alpha_0 = \rho(I - A_0^{-1} A_s) < 1$ and obtain an iteration for improving the convergence rate of Eq. (11.6). The matrix $\hat{G}^{-1}$ is not formed explicitly and the solution of systems involving $\hat{G}$ is achieved via the CG scheme, thus the only operations involved in the proposed nested iterative scheme (11.10) are matrix-vector multiplications and vector operations.

Our preconditioning strategy of the inner Richardson iteration is motivated by the study of Bank, Welfert and Yserentant [6] on a class of iterative methods for solving saddle-point problems. We extend it in this paper with some new results and a new analysis that relates the proposed iterative scheme to Uzawa's method. Further, we use our scheme for solving those *indefinite* linear systems that arise

from the mixed finite element discretization of 2D particulate flow problems, using P2-P1 type elements.

In what follows, we introduce the motivating application, the proposed nested iterative scheme, and analyze its convergence properties. We propose "optimal" approaches for the construction of $\hat{A}^{-1}$ and $\hat{G}^{-1}$ approximating A_s^{-1} and $(B^T \hat{A}^{-1} B)^{-1}$, or their actions on vectors, so as to assure convergence of our scheme. We also demonstrate the robustness of our nested iterative scheme as a preconditioner, its lack of sensitivity to changes in the fluid–particle systems, and its "scalability".

11.2 Motivating Application

Direct numerical simulation of particulate flows is of great value in a wide range of industrial applications such as enhancing productivity of oil reservoirs and the manufacturing process of polymers. From the numerical point of view, there are three classes of algorithms to handle such direct simulations, namely the space-time technique [35–38, 69], the "fictitious domain" formulation [30], and the "Arbitrary Lagrangian Eulerian" formulation, e.g. see [34, 42, 43, 48–51, 70]. All use finite elements for spatial discretization, and are based on a combined *weak* formulation, in which fluid and particle equations of motion are combined into a single *weak* equation of motion from which the hydrodynamic forces and torques on the particles have been eliminated, e.g. see [3, 4, 42] for details.

The particulate flow system is represented via the use of projection matrices that describe the constraints imposed on the system by the boundary conditions on the particle surfaces, where the vector velocity is reordered as $[\mathbf{u}_I^T, \mathbf{u}_\Gamma^T]^T$, in which $\mathbf{u}_\Gamma$ contains the components of the velocity field associated with the vertices on the particle boundaries, with the projection matrix applied to the decoupled system leading to a *nonsymmetric* (indefinite) saddle-point matrix with "borders".

For the direct numerical simulation of particulate flows, one must simultaneously integrate the Navier–Stokes equations, which govern the motion of the fluid, and the equations of rigid-body motion. These equations are coupled through the no-slip condition on the particle boundaries, and through the hydrodynamic forces and torques which appear in the equations of the rigid-body motion, e.g. see [3, 4, 42].

To establish a structurally symmetric matrix formulation of the coupled fluid–particle system, e.g. see [51] or [42], the first step is to assemble the matrices corresponding to the decoupled problem, where the no-slip condition is not taken into consideration. The Jacobian $\widetilde{J}$ of the decoupled fluid–particle system has the following algebraic form:

$$\widetilde{J} = \left[\begin{array}{cc|c} A & B & \\ B^T & 0 & \\ \hline & & M_p \end{array}\right],$$

in which case, the variable unknowns are ordered as follows

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{p} \\ \mathbf{U} \end{bmatrix} \quad \begin{array}{l} \mathbf{u}\text{: fluid velocity at each node} \\ \mathbf{p}\text{: fluid pressure} \\ \mathbf{U}\text{: particles velocity vector} \end{array}$$

and where M_p denotes the mass matrix of the n_p particles. M_p is block-diagonal and its size is $3n_p$ for $2D$ motion.

Since the approximate solution of the particulate flow problem is to be found in the subspace satisfying the no-slip condition, the constraints can be described in terms of a projection matrix. To clarify this further, the velocity unknowns may be divided into two categories, $\mathbf{u}_I$ for interior velocity unknowns and $\mathbf{u}_\Gamma$ for velocity unknowns on the surface of the particles. The Jacobian of the decoupled fluid–particle system is reordered accordingly, and hence the corresponding linear system is expressed in the following form:

$$\left[\begin{array}{ccc|c} A_{II} & A_{I\Gamma} & B_I & \\ A_{\Gamma I} & A_{\Gamma\Gamma} & B_\Gamma & \\ B_I^T & B_\Gamma^T & 0 & \\ \hline & & & M_p \end{array}\right] \begin{bmatrix} \mathbf{u}_I \\ \mathbf{u}_\Gamma \\ \mathbf{p} \\ \mathbf{U} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_I \\ \mathbf{f}_\Gamma \\ \mathbf{g} \\ \mathbf{f}_p \end{bmatrix}.$$

The no-slip condition on the surface of the particles requires that $\mathbf{u}_\Gamma = Q\mathbf{U}$, where Q is the projection matrix from the space of the surface unknowns onto the particle unknowns. Hence,

$$\begin{bmatrix} \mathbf{u}_I \\ \mathbf{u}_\Gamma \\ \mathbf{p} \\ \mathbf{U} \end{bmatrix} = \left[\begin{array}{ccc} I_{n\times n} & 0 & 0 \\ 0 & 0 & Q \\ 0 & I_{m\times m} & 0 \\ \hline 0 & 0 & I_{3n_p\times 3n_p} \end{array}\right] \begin{bmatrix} \mathbf{u}_I \\ \mathbf{p} \\ \mathbf{U} \end{bmatrix} = \widetilde{Q} \begin{bmatrix} \mathbf{u}_I \\ \mathbf{p} \\ \mathbf{U} \end{bmatrix}.$$

Finally, the Jacobian of the nonlinear coupled fluid–particle system can be written as $J = \widetilde{Q}^T \widetilde{J}\, \widetilde{Q}$, and we obtain the *nonsymmetric* bordered "saddle-point" problem,

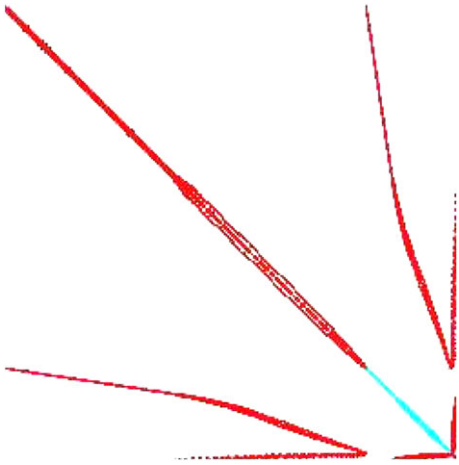
$$\left[\begin{array}{cc|c} A_{II} & B_I & A_{I\Gamma}Q \\ B_I^T & 0 & B_\Gamma^T Q \\ \hline Q^T A_{\Gamma I} & Q^T B_\Gamma & Q^T A_{\Gamma\Gamma} Q + M_p \end{array}\right] \begin{bmatrix} \mathbf{u}_I \\ \mathbf{p} \\ \mathbf{U} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_I \\ \mathbf{g} \\ \mathbf{f}_p \end{bmatrix}, \tag{11.11}$$

where the last block-column has a size equal to $3n_p$ for $2D$ motion.

Writing the Jacobian as

$$J = \left[\begin{array}{c|c} \mathscr{A} & \mathscr{B} \\ \hline \mathscr{C}^T & \mathscr{D} \end{array}\right], \tag{11.12}$$

Fig. 11.2 Field ordering



in which case, the variable unknowns are (always) ordered as follows, see Fig. 11.2

$$\begin{bmatrix} \mathbf{u}_I \\ \mathbf{p} \\ \mathbf{U} \end{bmatrix} \qquad \begin{array}{l} \mathbf{u}_I : \text{fluid velocity for the interior nodes} \\ \mathbf{p}: \ \text{fluid pressure} \\ \mathbf{U}: \ \text{particles velocity vector} \end{array}$$

and where the different block matrices are given by

$$\mathscr{A} = \begin{bmatrix} A_{II} & B_I \\ B_I^T & 0 \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}, \qquad \mathscr{B} = \begin{bmatrix} A_{I\Gamma} \\ B_\Gamma^T \end{bmatrix} Q \in \mathbb{R}^{(n+m)\times 3n_p},$$

$$\mathscr{C}^T = Q^T \begin{bmatrix} A_{\Gamma I} & B_\Gamma \end{bmatrix} \in \mathbb{R}^{3n_p\times(n+m)}, \qquad \mathscr{D} = Q^T A_{\Gamma\Gamma} Q + M_p \in \mathbb{R}^{3n_p\times 3n_p}.$$

It can be shown, e.g. see [3, 4], that

1. $\mathscr{A}$ and $\mathscr{D}$ are nonsingular, with $(\mathscr{D} + \mathscr{D}^T)/2$ symmetric positive definite,
2. $\mathscr{B}$ and $\mathscr{C}$ are of full-column rank,
3. $A_s = (A_{II} + A_{II}^T)/2$ is symmetric positive definite, and that A_{II} and $A_{\Gamma\Gamma}$ are positive stable, and
4. for the application considered here, Reynolds number ≤ 100, our choice of an effective time step, $\Delta t = 0.01$, and the discretization scheme adopted, the block diagonal matrix A_s is assured of having irreducibly diagonally dominant blocks.

Example 11.1 To verify numerically the above observations, we have conducted the simulation of a sedimentation experiment with 20 circular particles of diameter 1.0 in a channel of width 12.8 and length 124.0. Some information about the associated linear systems are displayed in Table 11.1, and the eigenvalue distribution of A_{II} is shown in Fig. 11.3. We clearly see that all the eigenvalues of A_{II} are on the right half of the complex plane, i.e., they all have positive real parts.

Table 11.1 Description of a small problem

Time	$\Delta t = 0.01$, $Re = 100.0$, Newton Iteration 5, 20 Particles				
	$\frac{1}{2}\|A_{II} + A_{II}^T\|_F$	$\frac{1}{2}\|A_{II} - A_{II}^T\|_F$	size(A_{II})	size($\mathscr{A}$)	cond($\mathscr{A}$)
$5\Delta t$	4×10^3	13	3994	4733	10^8

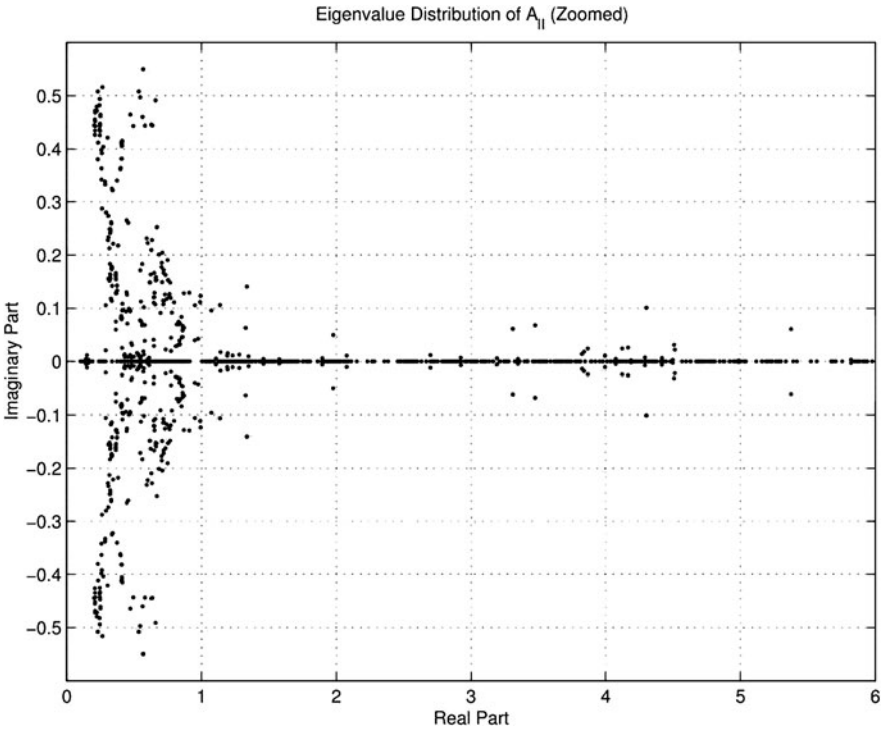


Fig. 11.3 Eigenvalue distribution of A_{II} at time step 5 (zoomed)

11.2.1 Properties of the Matrices

As the simulation time progresses, the structure of $\mathscr{A}$, its size and bandwidth vary, and its condition number increases. Generally, the flow simulation is characterized by three stages: the beginning, middle, and end of the simulation. Throughout the beginning and end stages, $\|A_s\|_F \gg \|A_{ss}\|_F$. In the middle stage, however, as the particulate flow becomes fully coupled, the Frobenius norm of the skew-symmetric part, $\|A_{ss}\|_F$, increases to approach $\|A_s\|_F$. Our experience indicates that Krylov subspace methods fail in solving Eq. (11.11) with classical ("black-box") preconditioners, even after only a few time steps, e.g., see [34, 42].

11.3 Solution Strategy

Since $\mathscr{D}$ is of much smaller dimension than $\mathscr{A}$ in Eq. (11.12), we solve Eq. (11.11) using the Schur complement approach by solving,

$$\left[\begin{array}{c|c} \mathscr{A} & \mathscr{B} \\ \hline & \mathscr{S}_1 \end{array}\right] \left[\begin{array}{c} \widetilde{\mathbf{x}} \\ \hline \widetilde{\mathbf{y}} \end{array}\right] = \left[\begin{array}{c} \mathbf{f} \\ \hline \hat{\mathbf{f}}_p \end{array}\right].$$

We first solve the Schur complement system $\mathscr{S}_1 \widetilde{\mathbf{y}} = \hat{\mathbf{f}}_p$ for $\widetilde{\mathbf{y}}$, where

$$\mathscr{S}_1 = \left[\mathscr{D} - \mathscr{C}^T \mathscr{A}^{-1} \mathscr{B}\right] \quad \text{and} \quad \hat{\mathbf{f}}_p = \left[\bar{\mathbf{f}}_p - \mathscr{C}^T \mathscr{A}^{-1} \mathbf{f}\right],$$

and once $\widetilde{\mathbf{y}}$ is obtained, $\widetilde{\mathbf{x}}$ is recovered by solving,

$$\mathscr{A} \widetilde{\mathbf{x}} = \mathbf{f} - \mathscr{B} \widetilde{\mathbf{y}},$$

via a preconditioned Krylov subspace method, such as GMRES [59]. In any case, the major task in solving Eq. (11.11) is the solution of a *nonsymmetric* saddle-point problem.

In the remainder of this paper, we concentrate on solving saddle-point systems of the form $\mathscr{A} \widetilde{\mathbf{x}} = \mathbf{b}$ using a Krylov subspace method such as GMRES, see Algorithm 11.1, with the indefinite preconditioner,

$$\mathscr{M} = \frac{1}{2}\left(\mathscr{A} + \mathscr{A}^T\right) = \begin{bmatrix} A_s & B_I \\ B_I^T & 0 \end{bmatrix}.$$

The inclusion of the exact representation of the (1,2) and (2,1) blocks of the preconditioner $\mathscr{M}$ leads one to hope for a more favorable distribution of the eigenvalues of the (left-)preconditioned linear system. The eigenvalues of the preconditioned coefficient matrix $\mathscr{M}^{-1}\mathscr{A}$ may be derived by considering the generalized eigenvalue problem

$$\begin{bmatrix} A_{II} & B_I \\ B_I^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \lambda \begin{bmatrix} A_s & B_I \\ B_I^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix},$$

which has an eigenvalue at 1 with multiplicity $2m$, and $(n-m)$ eigenvalues which are defined by the generalized eigenvalue problem, e.g. see [39]

$$Q_2^T A_{II}\, Q_2 \mathbf{z} = \lambda\, Q_2^T A_s\, Q_2 \mathbf{z}, \tag{11.13}$$

where $B_I = [Q_1 Q_2][\begin{smallmatrix} R \\ 0 \end{smallmatrix}]$. Thus, since $A_{II} = A_s + A_{ss}$, Eq. (11.13) is equivalent to

$$(\lambda - 1) = \frac{\mathbf{z}^\star\, Q_2^T A_{ss}\, Q_2 \mathbf{z}}{\mathbf{z}^\star\, Q_2^T A_s\, Q_2 \mathbf{z}}, \quad \forall \mathbf{z} \neq \mathbf{0},$$

Algorithm 11.1: Generalized Minimum RESidual (GMRES) [59]

1: Compute $\mathbf{r}_0 = \mathscr{M}^{-1}(\mathbf{b} - \mathscr{A}\,\widetilde{\mathbf{x}}_0), \beta = \| \mathbf{r}_0 \|_2,\ \mathbf{v}_1 = \mathbf{r}_0/\beta$.
2: **for** $j = 1, \ldots, m$ **do**
3: Compute $\mathbf{w} = \mathscr{M}^{-1}(\mathscr{A}\,\mathbf{v}_j)$.
4: **for** $i = 1, \ldots, j$ **do**
5: $h_{i,j} = \langle \mathbf{w}, \mathbf{v}_i \rangle$.
6: $\mathbf{w} = \mathbf{w} - h_{i,j}\,\mathbf{v}_i$.
7: **end for**
8: Compute $h_{j+1,j} = \| \mathbf{w} \|_2$;
and $\mathbf{v}_{j+1} = \mathbf{w}/h_{j+1,j}$.
9: **end for**
10: Define $V_m = \left[\mathbf{v}_1 \cdots \mathbf{v}_m\right],\ \bar{H}_m = \{h_{i,j}\}_{1 \le i \le j+1; 1 \le j \le m}$
11: Compute $\bar{\mathbf{y}}_m = \text{argmin}_{\bar{\mathbf{y}}} \| \beta\,\mathbf{e}_1 - \bar{H}_m\,\bar{\mathbf{y}} \|_2$;
and ;
$\widetilde{\mathbf{x}}_m = \widetilde{\mathbf{x}}_0 + V_m\,\bar{\mathbf{y}}_m$.
12: If satisfied Stop, else set $\widetilde{\mathbf{x}}_0 = \widetilde{\mathbf{x}}_m$;
and GOTO 1.

which means that if A_s is dominant, the eigenvalues λ are clustered around $(1 \pm i\gamma)$. Hence, the solution procedure is as follows. *Solve linear systems involving the matrix $\mathscr{A}$, which is the (1,1) (indefinite) saddle-point block of the Jacobian given in Eq. (11.11), by a preconditioned Krylov subspace method. Choosing an indefinite preconditioner $\mathscr{M}$ of the form Eq. (11.2), Step 3 of Algorithm 11.1, i.e., operations of the form $\mathbf{w} = \mathscr{M}^{-1}(\mathscr{A}\,\mathbf{v})$ are handled via the proposed nested iterative scheme.* Thus the algorithms presented in this paper are for solving the systems in the innermost loop of Fig. 11.1.

In what follows, we drop the subscript "I".

11.4 Proposed Nested Iterative Scheme

The matrix $\mathscr{M}$ can be factored as

$$\mathscr{M} = \begin{bmatrix} A_s & 0 \\ B^T & I \end{bmatrix} \begin{bmatrix} A_s^{-1} & 0 \\ 0 & -G \end{bmatrix} \begin{bmatrix} A_s & B \\ 0 & I \end{bmatrix}, \tag{11.14}$$

where $G = (B^T\ A_s^{-1}\ B)$. For many practical problems, an important feature of the system (11.14) is that the action of the matrices A_s^{-1} and G^{-1} can be approximated by "simple" matrices $\hat{A}^{-1}$ and $\hat{G}^{-1}$, in the sense that even though the computational cost of solving linear systems with the coefficient matrices $\hat{A}$ and $\hat{G}$ is low, the overall behavior of the algorithm lends itself to fast convergence. Other methods, with optimal order of computational complexity, are available for solving linear systems involving A_s, such as multigrid methods, e.g. see [5, 18, 71, 72].

The factorization (11.14) suggests an approximation of $\mathscr{M}$ given by

$$\hat{\mathscr{M}} = \begin{bmatrix} \hat{A} & 0 \\ B^T & I \end{bmatrix} \begin{bmatrix} \hat{A}^{-1} & 0 \\ 0 & -\hat{G} \end{bmatrix} \begin{bmatrix} \hat{A} & B \\ 0 & I \end{bmatrix}, \tag{11.15}$$

where $\hat{A}^{-1}$ and $\hat{G}^{-1}$ are approximations of A_s^{-1} and $(B^T\ \hat{A}^{-1}\ B)^{-1}$, respectively, and are assumed to be symmetric and positive definite.

Observing that the symmetric indefinite preconditioner $\hat{\mathscr{M}}$ in Eq. (11.15) is nonsingular, with exactly n positive and m negative eigenvalues, the nested iterative scheme for solving the symmetric saddle-point problem (11.3) consists of the preconditioned Richardson iteration (11.10),

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} + \begin{bmatrix} \hat{A} & B \\ B^T & -\hat{G} + B^T \hat{A}^{-1} B \end{bmatrix}^{-1} \left\{ \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix} - \begin{bmatrix} A_s & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} \right\},$$

where we consider the splitting $\mathscr{M} = \hat{\mathscr{M}} - \mathscr{N}$, with $\mathscr{N}$ being the defect matrix of the splitting.

This is equivalent to solving the following set of equations, which may be regarded as a version of a preconditioned inexact Uzawa algorithm with an additional correction step for $\mathbf{x}$, e.g. see [6, 74, 75]:

$$\begin{aligned} \hat{A}\left(\hat{\mathbf{x}}_{k+1} - \mathbf{x}_k\right) &= \mathbf{f} - [A_s\,\mathbf{x}_k + B\,\mathbf{y}_k], \\ \hat{G}\left(\mathbf{y}_{k+1} - \mathbf{y}_k\right) &= B^T \hat{\mathbf{x}}_{k+1} - \mathbf{g}, \\ \hat{A}\left(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}\right) &= -B\left(\mathbf{y}_{k+1} - \mathbf{y}_k\right). \end{aligned}$$

The corresponding algorithm is outlined by the following steps in Algorithm 11.2.

Algorithm 11.2: Nested iterative scheme

1: Initialize: $\mathbf{x} = \mathbf{x}_0$, $\mathbf{y} = \mathbf{y}_0$.
2: **for** $k = 0, 1, \ldots,$ until convergence **do**
3: Compute $\mathbf{r}_k = \mathbf{f} - \left[A_s\,\mathbf{x}_k + B\,\mathbf{y}_k\right]$.
4: Compute $\mathbf{s}_k = \mathbf{g} - B^T \mathbf{x}_k$.
5: Solve $\hat{A}\,\mathbf{c}_k = \mathbf{r}_k$.
6: Solve $\hat{G}\,\mathbf{d}_k = B^T \mathbf{c}_k - \mathbf{s}_k$.
7: Solve $\hat{A}\,\mathbf{c}_k = \mathbf{r}_k - B\,\mathbf{d}_k$.
8: Update $\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} + \begin{bmatrix} \mathbf{c}_k \\ \mathbf{d}_k \end{bmatrix}$.
9: **end for**

Remark 11.1 Step 7 in Algorithm 11.2 may be rearranged as

$$\mathbf{c}_k := \mathbf{c}_k - \hat{A}^{-1}(B\,\mathbf{d}_k),$$

where $\hat{A}^{-1}(B\,\mathbf{d}_k)$ is obtained as a byproduct of Step 6. This can save the application of $\hat{A}^{-1}$ at the end of every outer iteration, and thus improves the efficiency of the algorithm.

In each iteration step k of the nested iterative algorithm, five matrix-vector multiplications are required, namely $A_s\,\mathbf{x}_k$, $B\,\mathbf{y}_k$, $B^T\mathbf{x}_k$, $B^T\mathbf{c}_k$, and $B\,\mathbf{d}_k$; in addition to solving three systems, two of them involve $\hat{A}$ and the third involves $\hat{G}$.

11.5 Convergence Analysis of the Nested Iterative Scheme

The iteration matrix of the preconditioned Richardson iteration (11.10) is given by

$$\mathscr{K} = \left(\mathscr{I} - \hat{\mathscr{M}}^{-1}\,\mathscr{M}\right) = \hat{\mathscr{M}}^{-1}(\hat{\mathscr{M}} - \mathscr{M}),$$

where $\mathscr{M}$ and $\hat{\mathscr{M}}$ are given by Eqs. (11.2) and (11.9), respectively. Observing that

$$\hat{\mathscr{M}} - \mathscr{M} = \begin{bmatrix} \hat{A} - A_s & 0 \\ 0 & -\hat{G} + (B^T\hat{A}^{-1}B) \end{bmatrix},$$

and assuming that $\hat{A}$ and $\hat{G}$ are symmetric positive definite, then

$$\bar{\mathscr{K}} = \begin{bmatrix} \hat{A}^{\frac{1}{2}} & 0 \\ 0 & \hat{G}^{\frac{1}{2}} \end{bmatrix} \mathscr{K} \begin{bmatrix} \hat{A}^{-\frac{1}{2}} & 0 \\ 0 & \hat{G}^{-\frac{1}{2}} \end{bmatrix},$$

has the same eigenvalues as $\mathscr{K}$, and is given by

$$\bar{\mathscr{K}} = \begin{bmatrix} I - \bar{B}\bar{B}^T & \bar{B} \\ \bar{B}^T & -I \end{bmatrix} \begin{bmatrix} (I - \bar{A}) & 0 \\ 0 & -(I - \bar{G}) \end{bmatrix}, \tag{11.16}$$

in which

$$\bar{A} = \hat{A}^{-\frac{1}{2}} A_s\, \hat{A}^{-\frac{1}{2}} \qquad \in \mathbb{R}^{n\times n}, \tag{11.17}$$

$$\bar{B} = \hat{A}^{-\frac{1}{2}} B\, \hat{G}^{-\frac{1}{2}} \qquad \in \mathbb{R}^{n\times m}, \tag{11.18}$$

$$\bar{G} = \hat{G}^{-\frac{1}{2}}\left(B^T\hat{A}^{-1}B\right)\hat{G}^{-\frac{1}{2}} = \bar{B}^T\bar{B} \quad \in \mathbb{R}^{m\times m}. \tag{11.19}$$

Hence, the eigenvalues of $\mathscr{K}$ are close to zero when $\hat{A}^{-1}$ and $\hat{G}^{-1}$ are close to A_s^{-1} and $\left(B^T\hat{A}^{-1}B\right)^{-1}$, respectively.

Theorem 11.1 *Let α and β be the rates of convergence of the inner iterations* (11.6) *and* (11.8)*, respectively, defined by*

$$\alpha = \rho\left(I - \hat{A}^{-1}A_s\right) = \left\| I - \bar{A} \right\|_2 < 1,$$

and

$$\beta \;=\; \rho\big(\,I - \hat{G}^{-1}\big(B^T\,\hat{A}^{-1}B\big)\big) \;=\; \big\|\,I - \bar{G}\,\big\|_2 \;<\; 1,$$

then, in general, the iterative scheme (11.10) *is monotonically convergent if*

$$\max\{\alpha,\;\beta\} \;<\; \frac{\sqrt{5}-1}{2} \;\approx\; 0.6180.$$

Moreover, if $\beta \equiv 0$, *then a sufficient condition for convergence is* $\alpha < 1$, *and conversely, if* $\alpha \equiv 0$, *then it suffices to have* $\beta < 1$ *to guarantee convergence of Eq.* (11.10).

Proof We divide the proof into three cases.

Case 1: $\alpha \equiv 0$.
This special case corresponds to $(I - \bar{A}) \equiv 0$, i.e.,

$$\bar{\mathscr{K}} \;=\; \begin{bmatrix} 0 & -\bar{B}\,(I-\bar{G}) \\ 0 & (I-\bar{G}) \end{bmatrix} \quad\Longrightarrow\quad \rho(\bar{\mathscr{K}}) \;=\; \big\|\,I - \bar{G}\,\big\|_2 \;=\; \beta.$$

Case 2: $\beta \equiv 0$.
This special case corresponds to $(I - \bar{G}) \equiv 0$, i.e.,

$$\bar{\mathscr{K}} \;=\; \begin{bmatrix} (I-\bar{B}\,\bar{B}^T)(I-\bar{A}) & 0 \\ \bar{B}^T(I-\bar{A}) & 0 \end{bmatrix} \quad\Longrightarrow\quad \rho(\bar{\mathscr{K}}) \;=\; \rho\big[\big(I-\bar{B}\,\bar{B}^T\big)\big(I-\bar{A}\big)\big].$$

Therefore since $\alpha = \|\,I - \bar{A}\,\|_2$, $\rho(\bar{\mathscr{K}}) \le \alpha\,\|\,I - \bar{B}\,\bar{B}^T\,\|_2$.
Observing that in this case,

$$I - \bar{B}\,\bar{B}^T \;=\; I - \hat{A}^{-\frac{1}{2}}B\left(B^T\hat{A}^{-1}B\right)^{-1}B^T\hat{A}^{-\frac{1}{2}},$$

is an orthogonal projector, we have $\|\,I - \bar{B}\,\bar{B}^T\,\|_2 = 1$, and $\rho(\bar{\mathscr{K}}) \le \alpha$.
Case 3: This is the general case in which $\alpha,\ \beta < 1$. From Eq. (11.16), it is clear that

$$\begin{aligned} \rho\,(\bar{\mathscr{K}}) \;&\le\; \left\|\begin{bmatrix} (\,I-\bar{A}\,) & 0 \\ 0 & -\big(\,I-\bar{G}\,\big) \end{bmatrix}\right\|_2 \times \left\|\begin{bmatrix} I-\bar{B}\,\bar{B}^T & \bar{B} \\ \bar{B}^T & -I \end{bmatrix}\right\|_2, \\ &\le\; \max\{\,\alpha,\beta\,\}\left\|\begin{bmatrix} I-\bar{B}\,\bar{B}^T & \bar{B} \\ \bar{B}^T & -I \end{bmatrix}\right\|_2. \end{aligned}$$

Let the singular value decomposition of $\bar{B} = \hat{A}^{-\frac{1}{2}}B\,\hat{G}^{-\frac{1}{2}}$ be given by

$$\bar{B} \;=\; W\begin{bmatrix} \Omega \\ 0 \end{bmatrix}Y^T, \tag{11.20}$$

then

$$I - \bar{B}\,\bar{B}^T \;=\; W\begin{bmatrix} I_m - \Omega^2 & 0 \\ 0 & I_{n-m} \end{bmatrix}W^T,$$

where $W \in \mathbb{R}^{n\times n}$ and $Y \in \mathbb{R}^{m\times m}$ are orthogonal matrices and $\Omega \in \mathbb{R}^{m\times m}$ is a diagonal matrix containing the singular values ω_i of $\hat{A}^{-\frac{1}{2}} B \hat{G}^{-\frac{1}{2}}$ such that $1 > \omega_1 \geq \omega_2 \geq \cdots \geq \omega_m > 0$. Therefore,

$$\begin{bmatrix} I - \bar{B}\,\bar{B}^T & \bar{B} \\ \bar{B}^T & -I \end{bmatrix} = \begin{bmatrix} W & \\ & Y \end{bmatrix} \left[\begin{array}{cc|c} I_m - \Omega^2 & 0 & \Omega \\ 0 & I_{n-m} & 0 \\ \hline \Omega & 0 & -I_m \end{array}\right] \begin{bmatrix} W^T & \\ & Y^T \end{bmatrix},$$

and

$$\left\| \begin{bmatrix} I - \bar{B}\,\bar{B}^T & \bar{B} \\ \bar{B}^T & -I \end{bmatrix} \right\|_2 = \left\| \left[\begin{array}{cc|c} I_m - \Omega^2 & \Omega & \\ \Omega & -I_m & \\ \hline & & I_{n-m} \end{array}\right] \right\|_2 = \max\{1, \| T \|_2\},$$

in which T is the symmetric matrix

$$T = \begin{bmatrix} I_m - \Omega^2 & \Omega \\ \Omega & -I_m \end{bmatrix},$$

and where the eigenvalues of T are given by

$$\psi_i = -\frac{1}{2}\left[\omega_i^2 \pm \sqrt{\omega_i^4 + 4}\right], \quad i = 1, 2, \ldots, m.$$

Since $\omega_i^2 < 1$ due to the fact that $(I - \bar{G})$ is positive definite, then

$$\| T \|_2 = \frac{1}{2}\left[\omega_1^2 + \sqrt{\omega_1^4 + 4}\right] < \frac{1+\sqrt{5}}{2},$$

and

$$\rho(\bar{\mathscr{K}}) < \frac{1}{2}\left(1 + \sqrt{5}\right) \max\{\alpha, \beta\}.$$

Thus, to guarantee that $\rho(\bar{\mathscr{K}}) < 1$, it is sufficient to have

$$\frac{1}{2}\left(\sqrt{5}+1\right) \max\{\alpha, \beta\} < 1,$$

or

$$\max\{\alpha, \beta\} < \frac{1}{2}\left(\sqrt{5}-1\right) \approx 0.6180,$$

which completes the proof of Theorem 11.1. □

Remark 11.2 The previous results can be summarized as follows:

- If $\alpha = 0$ then $\rho(\mathscr{K}) = \beta$, and all eigenvalues $\lambda(\mathscr{K})$ are real.
- If $\beta = 0$ then $\rho(\mathscr{K}) \leq \alpha$, and all eigenvalues $\lambda(\mathscr{K})$ are real, as will be seen later in Lemma 11.2.

- Otherwise the eigenvalues $\lambda(\mathscr{K})$ are complex, and with the appropriate α-β relationship, $\rho(\mathscr{K}) < 1$.

Lemma 11.1 *Let* $\max\{\alpha, \beta\} < (\sqrt{5}-1)/2$, *and* $(I - \bar{A})$ *be positive definite. Then the eigenvalues of the iteration matrix* $\mathscr{K} = \mathscr{I} - \hat{\mathscr{M}}^{-1}\mathscr{M}$, *of the preconditioned Richardson iteration* (11.10), *lie to the right of the imaginary axis, i.e.,* $\Re(\lambda(\mathscr{K})) > 0$, *in addition to the fact that* $|\lambda(\mathscr{K})| < 1$.

Proof Consider the eigenvalue problem $\mathscr{K}\,\mathbf{v} = \lambda\,\mathbf{v}$ and let

$$\mathscr{D} = \begin{bmatrix} \hat{A}^{\frac{1}{2}} & 0 \\ 0 & \hat{G}^{\frac{1}{2}} \end{bmatrix},$$

then

$$\left(\mathscr{D}\,\mathscr{K}\,\mathscr{D}^{-1}\right)(\mathscr{D}\,\mathbf{v}) = \lambda\,(\mathscr{D}\,\mathbf{v}),$$

or

$$\begin{bmatrix} I - \bar{A} & 0 \\ 0 & I - \bar{G} \end{bmatrix}\begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} = \lambda \begin{bmatrix} I & \bar{B} \\ -\bar{B}^T & I - \bar{B}^T\bar{B} \end{bmatrix}\begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix},$$

where $\bar{A}$, $\bar{B}$, and $\bar{G}$ are as given in Eqs. (11.17)–(11.19). Using the singular value decomposition of $\bar{B}$ in Eq. (11.20), we get the eigenvalue problem

$$\left[\begin{array}{cc|c} I_m & 0 & \Omega \\ 0 & I_{n-m} & 0 \\ \hline -\Omega & 0 & I_m - \Omega^2 \end{array}\right]\mathbf{z} = \frac{1}{\lambda}\begin{bmatrix} \widetilde{A} & 0 \\ 0 & I - \Omega^2 \end{bmatrix}\mathbf{z},$$

where $\widetilde{A} = W^T(I - \hat{A}^{-\frac{1}{2}}A_s\hat{A}^{-\frac{1}{2}})W$ is symmetric positive definite. Since $\omega_i < 1$, we have

$$\tau \,\|\,\mathbf{z}\,\|_2^2 = \mathbf{z}^\star \begin{bmatrix} \widetilde{A} & 0 \\ 0 & I - \Omega^2 \end{bmatrix}\mathbf{z} > 0,$$

where $\mathbf{z}^\star = [\mathbf{z}_1^\star\ \mathbf{z}_2^\star\ \mathbf{z}_3^\star]$, and

$$\begin{aligned} \Re\left(\frac{1}{\lambda}\right) &= \frac{\|\,\mathbf{z}\,\|_2^2 - \mathbf{z}_3^\star\Omega^2\mathbf{z}_3}{\tau\,\|\,\mathbf{z}\,\|_2^2}, \\ &\geq \frac{1 - \omega_1^2}{\tau} > 0. \end{aligned}$$

Hence, $\Re(\lambda) > 0$, i.e., all the eigenvalues of $\mathscr{K}$ lie to the right of the imaginary axis, an ideal situation for acceleration via GMRES. □

Lemma 11.2 *For special case* 2, *when* $\hat{G} = (B^T \hat{A}^{-1} B)$, *i.e.*, $\beta = 0$, *the Richardson iteration matrix is given by*

$$\mathscr{K} = \begin{bmatrix} \hat{A}^{-\frac{1}{2}} & 0 \\ 0 & \hat{G}^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} M & 0 \\ N & 0 \end{bmatrix} \begin{bmatrix} \hat{A}^{\frac{1}{2}} & 0 \\ 0 & \hat{G}^{\frac{1}{2}} \end{bmatrix},$$

where

$$\begin{aligned} M &= (I - P)(I - \bar{A}) \quad \in \mathbb{R}^{n \times n}, \\ N &= \bar{B}^T (I - \bar{A}) \quad \in \mathbb{R}^{m \times n}, \end{aligned}$$

in which $\bar{A}$ *and* $\bar{B}$ *are as given in Eqs.* (11.17) *and* (11.18), *and* P *is the orthogonal projector*,

$$P = \bar{B}\bar{B}^T = \hat{A}^{-\frac{1}{2}} B \big(B^T \hat{A}^{-1} B\big)^{-1} B^T \hat{A}^{-\frac{1}{2}}.$$

Then $\mathscr{K}$ *has* $2m$ *zero eigenvalues*, *with* $\rho(\mathscr{K}) \leq \rho(I - \bar{A}) = \alpha < 1$, *and the submatrix of interest*,

$$\mathscr{K}_{11} = \hat{A}^{-\frac{1}{2}} (I - P)\big(I - \bar{A}\big)\hat{A}^{\frac{1}{2}},$$

has a complete set of eigenvectors X *with*

$$\kappa_2(X) \leq (1 + \hat{\mu}),$$

where $\kappa_2(\cdot)$ *denotes the spectral condition number*, *and*

$$\hat{\mu} < \left(\frac{1 + \sqrt{5}}{2}\right)\left(\frac{1}{\lambda^2_{\min}(C)}\right).$$

Here,

$$C = (I - P)(I - \bar{A})(I - P), \tag{11.21}$$

and $0 < |\lambda_{\min}(C)| = \min_i \{|\lambda_i(C)| \neq 0\}$.

Proof The fact that $\mathscr{K}$ has $2m$ zero eigenvalues is obvious from its structure and the fact that $(I - P)$ is an orthogonal projector of rank $(n - m)$. Also, from Theorem 11.1, we have $\rho(\mathscr{K}) \leq \|I - \bar{A}\|_2 = \alpha < 1$. Next, we consider the eigenvalue problem $\mathscr{K}_{11}\mathbf{z} = \lambda \mathbf{z}$, or

$$(I - P)\big(I - \bar{A}\big)\mathbf{w} = \lambda \mathbf{w}, \tag{11.22}$$

where $\mathbf{w} = \hat{A}^{\frac{1}{2}}\mathbf{z}$. Observing that the symmetric matrix C given by Eq. (11.21) has the same eigenvalues as Eq. (11.22) with an orthogonal set of eigenvectors $V = [V_1, V_2]$ such that

$$(I - P)\big(I - \bar{A}\big)(I - P)V_1 = V_1 \Lambda,$$

and

$$(I - P)\big(I - \bar{A}\big)(I - P)V_2 = 0,$$

in which $\Lambda = \mathrm{diag}(\lambda_i)$, with $\lambda_i \neq 0$, $i = 1, 2, \ldots, n - 2m$. Hence

$$(I - P)V_1 = V_1,$$

and

$$(I - P)V_2 = 0.$$

Since

$$V^T(I - P)(I - \bar{A})V = \begin{bmatrix} V_1^T(I - \bar{A})V_1 & V_1^T(I - \bar{A})V_2 \\ 0 & 0 \end{bmatrix},$$
$$= \begin{bmatrix} \Lambda & \hat{E} \\ 0 & 0 \end{bmatrix},$$

we can construct the nonsingular matrix

$$X = V \begin{bmatrix} I_{n-2m} & -\Lambda^{-1}\hat{E} \\ 0 & I_{2m} \end{bmatrix},$$

so that

$$X^{-1}(I - P)\big(I - \bar{A}\big)X = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}.$$

Consider the matrix

$$\hat{H} = \begin{bmatrix} I_{n-2m} & -H \\ 0 & I_{2m} \end{bmatrix},$$

where $H = \Lambda^{-1}\hat{E}$, then

$$\|\hat{H}\|_2^2 = 1 + \hat{\mu},$$

in which

$$\hat{\mu} = \Big[\mu^2 + \sqrt{\mu^4 + 4\mu^2}\Big] \Big/ 2,$$

with μ being the largest singular value of H. Consequently,

$$\|X\|_2^2 \leq (1 + \hat{\mu}).$$

This upper bound can be simplified further by observing that

$$\|H\|_2^2 \leq \frac{\alpha^2}{\lambda_{\min}^2(C)} < \frac{1}{\lambda_{\min}^2(C)},$$

in which $\alpha = \|I - \bar{A}\|_2$ and $\lambda_{\min}(C)$ is the smallest nonzero eigenvalue of C. Hence

$$\hat{\mu} < \left(\frac{1+\sqrt{5}}{2}\right)\left(\frac{1}{\lambda^2_{\min}(C)}\right),$$

and

$$\kappa_2(X) < 1 + \frac{1.618}{\lambda^2_{\min}(C)},$$

which completes the proof. □

11.6 Construction of $\hat{A}^{-1}$ and $\hat{G}^{-1}$

$\hat{A}^{-1}$ and $\hat{G}^{-1}$ are approximations of A_s^{-1} and $(B^T \hat{A}^{-1} B)^{-1}$, respectively. They are assumed to be symmetric and positive definite and are chosen such that $\alpha < 1$ and $\beta < 1$.

There are many ways to construct $\hat{A}^{-1}$ and $\hat{G}^{-1}$. For example, $\hat{A}$ can be taken as the incomplete Cholesky decomposition of A_s or other preconditioners of A_s. In this study, we always consider $\hat{A}$ and $\hat{G}$ corresponding to several iteration steps of a given iterative scheme for solving systems in Steps 5–7 of Algorithm 11.2. For example, suppose A_0 is a "simple" preconditioner for A_s, such that $\alpha_0 = \rho(I - A_0^{-1} A_s) < 1$, in which A_0 is obtained via a sparse approximate inverse scheme (SPAI), e.g. see [7, 33]. If we use the following "convergent" scheme, see Eq. (11.6), we have

$$\mathbf{x}_{k+1} = \left(I - A_0^{-1} A_s\right)\mathbf{x}_k + A_0^{-1}\underline{\mathbf{f}}, \quad k = 0, 1, 2, \tag{11.23}$$

for solving $A_s\,\mathbf{x} = \underline{\mathbf{f}}$. Choosing the initial iterate $\mathbf{x}_0 = \mathbf{0}$, we obtain

$$\mathbf{x}_3 = \left[\left(I - A_0^{-1} A_s\right)^2 + \left(I - A_0^{-1} A_s\right) + I\right] A_0^{-1}\underline{\mathbf{f}}.$$

Thus, we have *implicitly* generated $\hat{A}^{-1}$ as

$$\hat{A}^{-1} = \left[\left(I - A_0^{-1} A_s\right)^2 + \left(I - A_0^{-1} A_s\right) + I\right] A_0^{-1},$$

in which case it is easy to verify that

$$\rho\left(I - \hat{A}^{-1} A_s\right) = \rho^3\left(I - A_0^{-1} A_s\right) \ll 1.$$

Remark 11.3 One implicit matrix-vector multiplication with $\hat{A}^{-1}$ consists of 3 matrix-vector multiplications with A_0^{-1} and 2 matrix-vector multiplications with A_s. So the implicit acceleration via Eq. (11.23) results in more matrix-vector multiplications, but if A_0^{-1} is a diagonal matrix, for example, the additional cost is minimal, and the overall approach may be more economical than choosing a more accurate approximation of A_s^{-1}.

Remark 11.4 In solving the linear system in Step 6 of Algorithm 11.2, we use the conjugate gradient (CG) algorithm with $\hat{G} = B^T \hat{A}^{-1} B$, i.e., $\beta = 0$. Note the $B^T \hat{A}^{-1} B$ is never formed explicitly, and the major operation in each CG iteration is multiplying $\hat{G}$ by a vector. In our implementation in this study, however, we only solve the system in Step 6 approximately using a relaxed stopping criterion.

11.6.1 Construction of $\hat{A}^{-1}$

Recalling that each diagonal block of A_s is symmetric positive definite, and irreducibly diagonally dominant, one can construct a diagonal matrix A_0^{-1}, with positive elements such that $\rho(I - A_0^{-\frac{1}{2}} A_s A_0^{-\frac{1}{2}}) < 1$. Moreover, it can be easily verified that given such A_0^{-1}, then the action of the matrix $\hat{A}^{-1}$ can be *implicitly* generated via Eq. (11.23), such that $\rho(I - \hat{A}^{-\frac{1}{2}} A_s \hat{A}^{-\frac{1}{2}}) = \rho^3(I - A_0^{-\frac{1}{2}} A_s A_0^{-\frac{1}{2}}) = \alpha_0^3 \ll 1$.

Theorem 11.2 *Let the block diagonal matrix $A_s = [a_{ij}^{(s)}]$ be symmetric positive definite with each of its blocks irreducibly diagonally dominant, and let $A_0^{-1} = \mathrm{diag}(\delta_i)$ be the diagonal matrix that minimizes $\| I_n - A_s A_0^{-1} \|_F^2$. Then*

$$\delta_i = \frac{a_{ii}^{(s)}}{\| \mathbf{a}_i^{(s)} \|_2^2},$$

where $\mathbf{a}_i^{(s)}$ is the ith column of A_s, and the spectral radius

$$\rho\big(I - A_s A_0^{-1}\big) < 1.$$

Proof Let

$$\varphi = \| I_n - A_s A_0^{-1} \|_F^2 = \sum_{j=1}^{n} \| \mathbf{e}_j - \delta_j \mathbf{a}_j^{(s)} \|_2^2;$$

then φ is minimized when δ_j is chosen such that it minimizes $\| \mathbf{e}_j - \delta_j \mathbf{a}_j^{(s)} \|_2^2$, or

$$\delta_j = \frac{a_{jj}^{(s)}}{\| \mathbf{a}_j^{(s)} \|_2^2}.$$

The eigenvalue problem,

$$\big(I_n - A_s A_0^{-1}\big)\mathbf{u} = \lambda \mathbf{u},$$

yields

$$(1-\lambda) \;=\; \frac{\mathbf{v}^T A_s \mathbf{v}}{\mathbf{v}^T A_0 \mathbf{v}},$$

where $\mathbf{v} = A_0^{-1}\mathbf{u}$. Since both A_s and A_0 are symmetric positive definite, we have

$$(1-\lambda) \;>\; 0,$$

i.e., either λ is negative or has a positive value less than 1.

Next, consider the symmetric matrix $S = 2A_0 - A_s$, and let

$$\begin{aligned} D \;&=\; \operatorname{diag}\bigl(a_{11}^{(s)}, a_{22}^{(s)}, \ldots, a_{nn}^{(s)}\bigr), \\ A_s \;&=\; D - E, \end{aligned}$$

thus $D^{-\frac{1}{2}} A_s D^{-\frac{1}{2}} \;=\; I - D^{-\frac{1}{2}} E D^{-\frac{1}{2}}$, with $\rho(D^{-\frac{1}{2}} E D^{-\frac{1}{2}}) \;<\; 1$, e.g., see [53, Theorem 7.1.5, page 120].

Consequently,

$$D^{-\frac{1}{2}} S D^{-\frac{1}{2}} \;=\; \Theta \;+\; D^{-\frac{1}{2}} E D^{-\frac{1}{2}},$$

where

$$\Theta = \operatorname{diag}\left(2\frac{\|\mathbf{a}_j^{(s)}\|_2^2}{{a_{jj}^{(s)}}^2} - 1 \right) \;=\; \operatorname{diag}(1+\gamma_j) = I + \Gamma,$$

in which $\Gamma = \operatorname{diag}(\gamma_j)$, and $\gamma_j \;=\; \frac{1}{{a_{jj}^{(s)}}^2} \sum_{j \neq i} {a_{ij}^{(s)}}^2 \;>\; 0$.

Thus,

$$D^{-\frac{1}{2}} S D^{-\frac{1}{2}} \;=\; \Gamma + \bigl(I + D^{-\frac{1}{2}} E D^{-\frac{1}{2}}\bigr),$$

and since $\rho(D^{-\frac{1}{2}} E D^{-\frac{1}{2}}) < 1$, we see that S is symmetric positive definite. Moreover, it is easy to verify that

$$\begin{aligned} \mathbf{v}^T S \mathbf{v} \;&=\; 2\mathbf{v}^T A_0 \mathbf{v} - \mathbf{v}^T A_s \mathbf{v}, \\ &=\; \frac{(1+\lambda)}{(1-\lambda)}\, \mathbf{v}^T A_s \mathbf{v}, \end{aligned}$$

and as a result we have $-1 < \lambda < 1$, or $\rho(I - A_s A_0^{-1}) < 1$. □

In our numerical experiments, we consider also the more expensive approximate Cholesky factorization for obtaining $\hat{A} = R^T R \approx A_s$. In this case, we obtain the approximate factorization using a numerical drop tolerance as well as a prescribed maximum fill-in per row.

Table 11.2 shows a few problem instances in the particulate flow simulation. Results in Table 11.3 show that the reduction in the number of inner iterations

Table 11.2 Description of the set of problems

Time	$\Delta t = 0.01$, $Re = 100.0$, Newton Iteration 5, 20 Particles				
	$\frac{1}{2}\|A_{II} + A_{II}^T\|_F$	$\frac{1}{2}\|A_{II} - A_{II}^T\|_F$	size(A_{II})	size($\mathscr{A}$)	cond($\mathscr{A}$)
$5\Delta t$	4×10^3	13	3994	4733	10^8
$10\Delta t$	4×10^3	8	4336	5213	10^8
$20\Delta t$	4×10^3	15	4179	4920	10^8

Table 11.3 Results with SPAI(A_s, diag) vs. IC(A_s, 15, 1.0e-3)

GMRES(20)							
Problem	SPAI(A_s, diag)				IC(A_s, 15, 1.0e-3)		
Instance	α_0	$\alpha = \alpha_0^3$	inner	outer	α	inner	outer
$5\,\Delta t$	0.8748	0.6694	12	1	0.4912	8	1
$10\,\Delta t$	0.8842	0.6913	11	1	0.5217	8	1
$20\,\Delta t$	0.8617	0.6398	12	1	0.5021	8	1

(Richardson iterations in Algorithm 11.2) realized by using the expensive explicit generation of $\hat{A}$ via the approximate Cholesky factorization, is not sufficient to justify its use. Note that the number of inner iterations listed in Table 11.3 represents the number of iterations needed for a single call of Algorithm 11.2. Thus, for example, for the problem arising at time $5\Delta t$, with 1 outer iteration of GMRES(20), the total number of inner iterations is 240, which is still much more economical than solving systems of the form Eq. (11.3) directly within GMRES; see also Sect. 11.7.1.

11.6.2 Implicit Generation of Variable $\hat{G}_k^{-1}$

In Step 6 of Algorithm 11.2, we need to solve linear systems of the form

$$\hat{G}\mathbf{d}_k = \mathbf{h}_k,$$

to determine an approximate solution $\hat{\mathbf{d}}$ via the conjugate gradient (CG) scheme where we replace $\hat{G}$ by $(B^T\hat{A}^{-1}B)$. Thus, the approximation of the action of $(B^T\hat{A}^{-1}B)^{-1}$ on $\mathbf{h}_k$ varies in each CG iteration.

The following theorem gives an explanation as to why there is no need to solve the inner system (Step 6) accurately, i.e., at each CG iteration j, there is $\hat{G}_j$ such that $\hat{G}_j\hat{\mathbf{d}} = (B^T\hat{A}^{-1}B)\,\mathbf{d}$, where $\hat{\mathbf{d}}$ is close to $\mathbf{d}$ in the $(B^T\hat{A}^{-1}B)$-norm defined by

$$\|\mathbf{y}\|^2_{(B^T\hat{A}^{-1}B)} = \left\langle \mathbf{y},\, \left(B^T\hat{A}^{-1}B\right)\mathbf{y} \right\rangle, \quad \forall \mathbf{y} \in \mathbb{R}^m,$$

where $\langle\cdot,\,\cdot\rangle$ is the usual Euclidean inner-product.

Theorem 11.3 (Bank, Welfert and Yserentant [6]) *Let* $(B^T \hat{A}^{-1} B)$ *be a symmetric and positive definite* $m \times m$ *matrix, and let* $\mathbf{d}, \hat{\mathbf{d}} \in \mathbb{R}^m$ *satisfy*

$$\| \mathbf{d} - \hat{\mathbf{d}} \|_{(B^T \hat{A}^{-1} B)} \leq \beta \| \mathbf{d} \|_{(B^T \hat{A}^{-1} B)},$$

with $0 \leq \beta < 1$. *Then there exists a symmetric positive definite matrix* $\hat{G}$ *with*

$$\hat{G}\hat{\mathbf{d}} = \left(B^T \hat{A}^{-1} B\right)\mathbf{d},$$

and

$$\| I - \hat{G}^{-\frac{1}{2}} \left(B^T \hat{A}^{-1} B\right) \hat{G}^{-\frac{1}{2}} \|_2 \leq \beta.$$

Proof The proof is by construction and can be found in [6]. □

Let each CG iteration j yield an approximate solution $\mathbf{d}_{k,j}$ with residual $\mathbf{r}_{k,j}$ given by

$$\begin{aligned} \mathbf{r}_{k,j} &= \mathbf{h}_k - \left(B^T \hat{A}^{-1} B\right)\mathbf{d}_{k,j}, \\ &= \mathbf{h}_k - \left(B^T \hat{A}^{-1} B\right)\hat{G}_j^{-1}\mathbf{h}_k, \\ &= \left[I - \left(B^T \hat{A}^{-1} B\right)\hat{G}_j^{-1}\right]\mathbf{h}_k. \end{aligned}$$

Therefore,

$$\left\| I - \hat{G}_j^{-\frac{1}{2}} \left(B^T \hat{A}^{-1} B\right) \hat{G}_j^{-\frac{1}{2}} \right\|_2 \geq \frac{\| \mathbf{r}_{k,j} \|_2}{\| \mathbf{h}_k \|}.$$

In general, there exists a $\widehat{\gamma} \approx 1$, e.g. see [68, p. 194], such that

$$\left\| I - \hat{G}_j^{-\frac{1}{2}} \left(B^T \hat{A}^{-1} B\right) \hat{G}_j^{-\frac{1}{2}} \right\|_2 \approx \widehat{\gamma} \frac{\| \mathbf{r}_{k,j} \|_2}{\| \mathbf{h}_{k,j} \|_2}.$$

Consequently, choosing the stopping criterion $\| \mathbf{r}_{k,j} \|_2 / \| \mathbf{h}_k \|_2 \leq 10^{-2}$ will almost guarantee a value of $\beta = \mathcal{O}(10^{-2})$.

In order to verify this last observation, we conducted a set of numerical experiments in which we solve Eq. (11.3) using the preconditioned Richardson iteration (11.10) using Algorithm 11.2 with a relative residual stopping criterion of 10^{-6}. Here, the system in Step 6 is solved using the conjugate gradient algorithm (without preconditioning) with $\hat{G}$ replaced by $(B^T \hat{A}^{-1} B)$. Table 11.4 shows the results for varying levels of the CG relative residuals stopping criterion (`tol_CG`), from 10^{-6} to 10^{-1}, for a sample problem.

The vectors $\mathbf{b} = [\mathbf{f}^T \mathbf{g}^T]^T$, $\widetilde{\mathbf{r}}_k = [\mathbf{r}_k^T \ \mathbf{s}_k^T]^T$, and $\mathbf{w}_k = [\mathbf{x}_k^T \ \mathbf{y}_k^T]^T$ are as given in Algorithm 11.2, and $\boldsymbol{\delta}\mathbf{w}_k = \mathbf{w}_\star - \mathbf{w}_k$, in which $\mathbf{w}_\star$ is the exact solution of (11.3). It is clear that using a `tol_CG` of 10^{-2} produces just as satisfactory a result had we used a `tol_CG` of 10^{-6}. This result confirms Theorem 11.3.

Table 11.4 Inner-inner iteration: CG method in Step 6 of Algorithm 11.2

Problem Instance: $t = 20\,\Delta t$			
tol_CG	Richardson iters	$\|\tilde{\mathbf{r}}_k\|_2/\|\tilde{\mathbf{r}}_0\|_2$	$\|\boldsymbol{\delta}\mathbf{w}_k\|_2/\|\mathbf{b}\|_2$
1.0e-6	12	8.4×10^{-7}	10^{-4}
1.0e-5	12	8.4×10^{-7}	10^{-4}
1.0e-4	12	8.4×10^{-7}	10^{-4}
1.0e-3	12	8.4×10^{-7}	10^{-4}
1.0e-2	12	8.4×10^{-7}	10^{-4}
1.0e-1	8	4.7×10^{-7}	10^{-3}

Table 11.5 Values of α

Problem Instance	SPAI-0		IC
	α_0	$\alpha = \alpha_0^3$	α
$5\,\Delta t$	0.8748	0.6694	0.4912
$10\,\Delta t$	0.8842	0.6913	0.5217
$20\,\Delta t$	0.8617	0.6398	0.5021

Table 11.6 Inner-outer iterations

$\hat{G} = (B^T \hat{A}^{-1} B) \implies \beta \approx 0$				
Problem Instance	GMRES(20)			
	SPAI(A_s, diag)		IC(A_s, 15, 1.0e−4)	
	Richardson iters	outer	Richardson iters	outer
$5\Delta t$	12	1	8	1
$10\Delta t$	11	1	8	1
$20\Delta t$	12	1	8	1

In what follows, we generate A_0^{-1} via SPAI-0, and then obtain *implicitly* the action of $\hat{A}^{-1}$ on a vector via Eq. (11.23). For a sample problem at different time Steps, Table 11.5 shows the spectral radii $\alpha = \rho(I - \hat{A}^{-1} A_s)$ as well as those when $\hat{A} = R^T R$ in which R^T is the approximate Cholesky factor of A_s. Also, for solving systems in Step 6 of Algorithm 11.2, $\hat{G}\mathbf{d} = \mathbf{h}$, we use the conjugate gradient scheme with a relaxed stopping criterion, in which $\hat{G} = (B^T \hat{A}^{-1} B)$.

Now, using our complete nested iterative scheme, illustrated in Fig. 11.1, on the same set of sample problems, with GMRES(20), yields a solution satisfying the outer iteration stopping criterion of a relative residual less than or equal to 10^{-6}, only after one outer GMRES iteration, see Table 11.6.

Again, we show that using an approximate Cholesky factorization to generate $\hat{A}$ does not reduce the number of inner (Richardson) iterations sufficiently to justify the additional cost in each time step.

Table 11.7 Robustness of the nested iterative scheme

$\hat{G} = (B^T \hat{A}^{-1} B) \implies \beta \approx 0$: CG method with $\|\tilde{\mathbf{r}}_k\|_2 / \|\tilde{\mathbf{r}}_0\|_2 \leq 10^{-3}$								
t	n_p	$(n+m)$	$\hat{A}$	$\alpha = \alpha_0^3$	GMRES(k)			
					inner	outer	k	ARMS
$20\,\Delta t$	20	8777	SPAI(A_s, diag)	0.6913	13	1	20	10
$100\,\Delta t$	240	95749	IC(A_s, 15, 10^{-4})	0.6514	10	2	50	†
			SPAI(A_s, diag)	0.7216	14	3	50	†
$200\,\Delta t$	240	111326	IC(A_s, 15, 10^{-4})	0.6911	12	3	50	†
			SPAI(A_s, diag)	0.7502	15	4	50	†

Table 11.8 Parameters for ARMS

bsize	nlev	fill_{I}	$\text{fill}_{\text{last}}$	$\text{fill}_{\text{ILUT}}$	$\text{droptol}_{\text{I}}$	$\text{droptol}_{\text{last}}$
500–1000	2–5	60	50	50	0.0001	0.001

11.7 Numerical Experiments

We show the robustness of our nested iterative scheme illustrated in Fig. 11.1 by solving linear systems (11.1) for varying sizes, as the number of particles increases and the mesh size decreases, and at different time steps from $10\Delta t$ to $200\Delta t$, as the dominance of A_s (vs. A_{ss}) decreases.

Adopting a stopping criterion of a 10^{-6} relative residual for the outer GMRES iterations, our results are shown in Tables 11.7, 11.8 and 11.10. In Table 11.7 we give the number of inner (Richardson) and outer (GMRES) iterations. We also note that for $\hat{A}^{-1}$ generated via SPAI-0, with the implicit acceleration (11.23), and the systems $(B^T \hat{A}^{-1} B)\,\mathbf{d} = \mathbf{h}$ solved via the CG scheme with a relaxed stopping criterion, the scheme is remarkably robust succeeding in solving all the linear systems arising in the particulate flow simulations of Newtonian fluids.

In Table 11.7, we compare our nested iterative scheme, and GMRES with Saad's "black-box" preconditioner, "Algebraic Recursive Multilevel Solver", see [60], applied to Eq. (11.1). In using ARMS(*nlev*), we employ $nlev = 2$ levels and in case of failure we increase *nlev* to 5. Table 11.8 shows the parameters that need to be set up for ARMS. These parameters can be fine-tuned for a particular system to assure success. In Table 11.9, we compare our scheme with GMRES preconditioned via the ILUT factorization of $\mathscr{A}$. Note that both general purpose preconditioners, ILUT and ARMS, could fail for our saddle-point problems. Finally, in Table 11.10 we illustrate "scalability" of our nested iterative scheme in the sense that the number of inner iterations in any given pass of Algorithm 11.2 remains *almost* constant with only a modest increase in the number of outer (GMRES) iterations.

Table 11.9 Comparison with ILUT

Size($\mathscr{A}$)	GMRES(k)								
	nested scheme					ILUT(p,τ)			
	$\hat{A}$	$\alpha = \alpha_0^3$	k	inner	outer	p	τ	iters	k
29816	IC(A_s, 15, 10^{-3})	0.5198	20	8	2	15	0.0001	89	20
65471	IC(A_s, 15, 10^{-3})	10^{-4}	20	2	2	15	0.0001	42	20
80945	IC(A_s, 15, 10^{-3})	10^{-4}	50	2	2	15	0.0001	–	100
95749	SPAI(A_s, diag)	0.7216	50	14	3	15	0.0001	–	100
111326	SPAI(A_s, diag)	0.7502	50	15	4	15	0.0001	–	100

Table 11.10 "Scalability" of nested iterative scheme

$n_p = 20$, $t = 70\,\Delta t$, $\hat{A}^{-1} = \texttt{SPAI}(A_s, \texttt{diag})$				
$(n+m)$	$\alpha = \alpha_0^3$	GMRES(k)		
		inner	outer	k
3872	0.6782	12	1	20
6157	0.6973	12	1	20
10217	0.7012	13	2	20
31786	0.7314	14	2	20
56739	0.7196	14	3	20
81206	0.7512	15	3	40
105213	0.7419	15	3	40

11.7.1 Comparison with Other Preconditioners

We compare our scheme with algorithms of the form displayed in Fig. 11.4, i.e., GMRES inner–outer iterations. Here we consider two preconditioners

$$\mathscr{M}_1 = \begin{bmatrix} A & 0 \\ 0 & B^T A^{-1} B \end{bmatrix}, \quad \text{and} \quad \mathscr{M}_2 = \begin{bmatrix} A_s & B \\ B^T & 0 \end{bmatrix}.$$

In the block-diagonal preconditioner case, the preconditioned matrix $\mathscr{P}_1 = \mathscr{M}_1^{-1}\mathscr{A}$ has at most four distinct eigenvalues [52], namely 0, 1, and $(1 \pm \sqrt{5})/2$. Thus, it directly follows that for any vector, the associated Krylov subspace is of dimension at most three if $\mathscr{P}_1$ is nonsingular (or four if $\mathscr{P}_1$ is singular). Thus, any Krylov subspace iterative method with an optimality property, such as GMRES, will terminate in at most three iterations in *exact arithmetic*. As for the indefinite preconditioner $\mathscr{M}_2$, more favorable distribution of the eigenvalues of the (left-)preconditioned linear system $\mathscr{P}_2 = \mathscr{M}_2^{-1}\mathscr{A}$, is expected. Since solving with the Schur complement $(B^T A^{-1} B)$ is too expensive, the block-diagonal precondi-

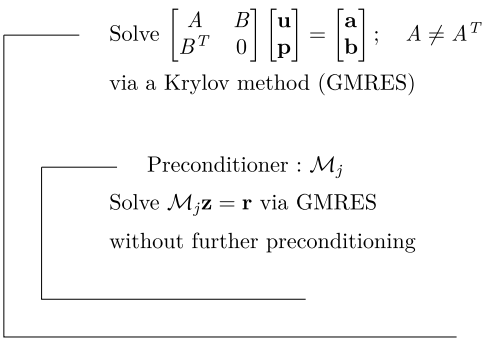


Fig. 11.4 A preconditioned Krylov method

tioner $\mathscr{M}_1$ is approximated by

$$\widetilde{\mathscr{M}_1} = \begin{bmatrix} \text{diag}(A) & 0 \\ 0 & B^T(\text{diag}(A))^{-1}B \end{bmatrix},$$

and

$$\widetilde{\mathscr{M}_2} = \begin{bmatrix} A & 0 \\ 0 & B^T(\text{diag}(A))^{-1}B \end{bmatrix}.$$

For the sake of completeness, we have added another popular symmetric indefinite preconditioner considered by Perugia and Simoncini in [55] for the solution of the stabilized symmetric saddle point problem that arises in mixed finite element approximations of magnetostatic problems,

$$\mathscr{M}_3 = \begin{bmatrix} I & B \\ B^T & 0 \end{bmatrix},$$

which is essentially a special case of Eq. (11.9) with $\hat{A} = I$ and $\hat{G} = B^T B$.

To evaluate the performance of the different preconditioners used in conjunction with inner–outer GMRES, e.g. see [58, 66, 67], numerical experiments have been performed on the set of problems displayed in Table 11.2. The results presented in Table 11.11 show that preconditioners $\widetilde{\mathscr{M}_1}$, $\widetilde{\mathscr{M}_2}$ and $\mathscr{M}_3$ require more than one outer iteration, whereas $\mathscr{M}_2$ seems to be the most effective competitor to our nested iterative scheme. Timing experiments on a uniprocessor, however, show that our nested scheme is at least 8 times faster than the other preconditioners shown in Table 11.11, used in the inner–outer GMRES setting of Fig. 11.4.

11.7.2 The Driven-Cavity Steady-State Case

Finally, we have used our nested scheme for obtaining the steady state solution of the Navier–Stokes equations modeling the incompressible fluid flow within a "leaky" two-dimensional lid-driven cavity problem in a square domain $-1 \leq x,\ y \leq$

Table 11.11 Performance of GMRES(20) with the different preconditioners

Time instance	$\Delta t = 0.01$, $Re = 100.0$, Newton Iteration 5				
	inner–outer iterations	Block-diagonal		Indefinite	
		$\widetilde{\mathscr{M}_1}$	$\widetilde{\mathscr{M}_2}$	$\mathscr{M}_2$	$\mathscr{M}_3$
$5\Delta t$	inner iters	10	6	7	14
$n_{\mathscr{A}} = 4733$	outer iters	8	2	1	11
$10\Delta t$	inner iters	14	8	15	20
$n_{\mathscr{A}} = 5213$	outer iters	8	2	1	12
$20\Delta t$	inner iters	12	7	10	16
$n_{\mathscr{A}} = 4920$	outer iters	8	2	1	11

Table 11.12 outer iteration: GMRES(k); a direct method for solving (11.3)

GMRES(10)			GMRES(20)		
mesh	outer	$\|$ residual $\|_2$	mesh	outer	$\|$ residual $\|_2$
8×8	3	1×10^{-6}	8×8	2	5×10^{-10}
16×16	4	2×10^{-7}	16×16	2	2×10^{-8}
32×32	4	1×10^{-7}	32×32	2	2×10^{-8}

1 with fluid viscosity of 0.01. The boundary condition for this model problem is $\mathbf{u}_x = \mathbf{u}_y = 0$ on the three walls ($x, y = -1;\ x = 1$), and $\mathbf{u}_x = 1$, $\mathbf{u}_y = 0$ on the moving wall ($y = 1$). Using Picard's iteration and mixed finite element ($Q2/Q1$) approximation of the resulting linearized equations (Oseen problems), we obtain systems of the form (11.1), derived from Picard's ninth iteration. Moreover, we see that even though the (1,1) block in Eq. (11.1) no longer has the advantage of the term [$(1/\Delta t) \times$ mass matrix], all the properties outlined above of its symmetric part still hold. Using a uniform mesh, the tables below show the effectiveness of our nested iterative scheme and its independence of the mesh size.

In Tables 11.12–11.13 we give the number of outer iterations of GMRES(10) and GMRES(20) needed to reach a residual of 2-norm less than or equal to 10^{-6} for solving Eq. (11.1). In Tables 11.12, similar to the Golub–Wathen study [31], we solve systems involving the preconditioner (11.2) using a direct scheme.

In Tables 11.13, we present similar results, except that we use our nested scheme shown in Fig. 11.1, with Algorithm 11.2 limited to only four iterations. The results shown in Tables 11.13 are the same whether the linear system in Step 6 of Algorithm 11.2 is solved directly, or solved using the conjugate gradient scheme with a relative residual stopping criterion of 10^{-2}. For much larger problems, however, the cost of direct solvers will be much higher than the CG scheme with a relaxed stopping criterion. Furthermore, the 2-norm of the residuals in Tables 11.12 and 11.13 are essentially the same. This demonstrates the effectiveness of our nested

Table 11.13 outer iteration: GMRES(k); the nested iterative scheme for solving (11.3)

GMRES(10)				GMRES(20)			
mesh	outer	$\|\text{residual}\|_2$	inner	mesh	outer	$\|\text{residual}\|_2$	inner
8×8	3	9×10^{-7}	4	8×8	2	5×10^{-10}	4
16×16	4	4×10^{-7}	4	16×16	2	7×10^{-8}	4
32×32	8	2×10^{-7}	4	32×32	4	5×10^{-8}	4

iterative scheme, not only for obtaining time-accurate solutions of the particulate flow problems, but also for the steady-state for driven cavity problem outlined here.

Finally, we would like to state that using flexible-type GMRES, to allow for changes in the preconditioner from one outer iteration to another, has resulted in inferior performance compared to that reported in Table 11.13. We would like also to mention that GMRES(20), without preconditioning, requires 592 iterations for the 8×8 mesh, and failed to achieve a residual of 2-norm $\leq 10^{-5}$ after 2000 iterations, for the 16×16 and 32×32 meshes.

11.8 Conclusion

We have presented a "nested iterative scheme" for solving saddle-point problems which can be regarded as a preconditioned inexact Uzawa algorithm with an additional correction step. The algorithm is essentially a preconditioned Krylov subspace method in which the preconditioner is itself a saddle-point problem. We propose a preconditioned Richardson iteration, with monotone convergence, for handling those inner iterations, i.e. for solving those systems involving the preconditioner. It should be noted that this Richardson scheme can be very effective in solving symmetric saddle-point problems in which the (1,1) block is symmetric positive definite.

We have used our nested iterative scheme for solving those nonsymmetric saddle-point problems that arise from the mixed finite element discretization of particulate flows, in which the fluid is incompressible. We have shown that an "inexpensive" preconditioner can be easily constructed, i.e. by constructing $\hat{A}^{-1}$ and $\hat{G}^{-1}$. In particular, we have shown that it is sufficient to have $\hat{A}^{-1} = \text{SPAI}(A_s, \text{diag})$, accelerated implicitly by three iterations, and to have the action of $\hat{G}^{-1}$ a close approximation of the action of $(B^T \hat{A}^{-1} B)^{-1}$. This latter implicit construction of $\hat{G}^{-1}$ is accomplished by solving systems involving $(B^T \hat{A}^{-1} B)$ via the Conjugate Gradient method with a relaxed stopping criterion.

We have compared our solution strategy of systems involving the adopted preconditioner with other preconditioners available in the literature. The resulting nested scheme proved to be more robust and more economical than others for handling those particulate flow simulations. Moreover, our scheme proved to be "scalable", and insensitive to changes in the fluid–particle system.

Finally, we should point out that all basic operations of our nested iterative scheme are amenable to efficient implementation on parallel computers.

Acknowledgements This work has been done in collaboration with Prof. Ahmed Sameh, and the author would like to acknowledge him for his continuous support.

References

1. Arrow, K., Hurwicz, L., Uzawa, H.: Studies in Nonlinear Programming. Stanford University Press, Stanford (1958)
2. Babuska, I.: The finite element method with Lagrangian multipliers. Numer. Math. **20**, 179–192 (1973)
3. Baggag, A.: Linear system solvers in particulate flows. Ph.D. thesis, Department of Computer Science, University of Minnesota (2003)
4. Baggag, A., Sameh, A.: A nested iterative scheme for indefinite linear systems in particulate flows. Comput. Methods Appl. Mech. Eng. **193**, 1923–1957 (2004)
5. Bank, R.E., Dupont, T., Yserentant, H.: The hierarchical basis multigrid method. Numer. Math. **52**, 427–458 (1988)
6. Bank, R.E., Welfert, B.D., Yserentant, H.: A class of iterative methods for solving saddle point problems. Numer. Math. **55**, 645–666 (1990)
7. Barnard, S., Grote, M.: A block version of the SPAI preconditioner. In: Hendrickson, B., Yelick, K., Bishof, C. (eds.) Proceedings of the Ninth SIAM Conference on Parallel Processing for Scientific Computing, March 22–24. SIAM, Philadelphia (1999). CD-ROM
8. Benzi, M., Golub, G.H.: An iterative method for generalized saddle point problems. SIAM J. Matrix Anal. (2012, to appear)
9. Braess, D., Sarazin, R.: An efficient smoother for the stokes problem. Appl. Numer. Math. **23**, 3–19 (1997)
10. Bramble, J.H., Leyk, Z., Pasciak, J.E.: Iterative schemes for non-symmetric and indefinite elliptic boundary value problems. Math. Comput. **60**, 1–22 (1993)
11. Bramble, J.H., Pasciak, J.E.: A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. Math. Comput. **50**, 1–18 (1988)
12. Bramble, J.H., Pasciak, J.E.: Iterative techniques for time dependent Stokes problems. Comput. Math. Appl. **33**, 13–30 (1997)
13. Bramble, J.H., Pasciak, J.E., Vassilev, A.T.: Analysis of the inexact Uzawa algorithm for saddle point problems. SIAM J. Numer. Anal. **34**, 1072–1092 (1997)
14. Bramble, J.H., Pasciak, J.E., Vassilev, A.T.: Uzawa type algorithms for nonsymmetric saddle point problems. Math. Comput. **69**, 667–689 (2000)
15. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer, New York (1991). ISBN 3-540-97582-9
16. Dyn, N., Ferguson, W.: The numerical solution of equality-constrained quadratic programming problems. Math. Comput. **41**, 165–170 (1983)
17. Elman, H., Silvester, D.: Fast nonsymmetric iterations and preconditioning for Navier–Stokes equations. SIAM J. Sci. Comput. **17**, 33–46 (1996)
18. Elman, H.C.: Multigrid and Krylov subspace methods for the discrete Stokes equations. Tech. Rep. 3302, Institute for Advanced Computer Studies (1994)
19. Elman, H.C.: Perturbation of eigenvalues of preconditioned Navier–Stokes operators. SIAM J. Matrix Anal. Appl. **18**, 733–751 (1997)
20. Elman, H.C.: Preconditioning for the steady-state Navier–Stokes equations with low viscosity. SIAM J. Sci. Comput. **20**, 1299–1316 (1999)
21. Elman, H.C.: Preconditioners for saddle point problems arising in computational fluid dynamics. Appl. Numer. Math. **43**, 75–89 (2002)

22. Elman, H.C., Golub, G.H.: Inexact and preconditioned Uzawa algorithms for saddle point problems. SIAM J. Numer. Anal. **31**, 1645–1661 (1994)
23. Elman, H.C., Silvester, D.J., Wathen, A.J.: Iterative methods for problems in computational fluid dynamics. In: Chan, R., Chan, T., Golub, G. (eds.) Iterative Methods in Scientific Computing. Springer, Singapore (1997)
24. Elman, H.C., Silvester, D.J., Wathen, A.J.: Performance and analysis of saddle point preconditioners for the discrete steady-state Navier–Stokes equations. Numer. Math. **90**, 641–664 (2002)
25. Falk, R.: An analysis of the finite element method using Lagrange multipliers for the stationary Stokes equations. Math. Comput. **30**, 241–269 (1976)
26. Falk, R., Osborn, J.: Error estimates for mixed methods. RAIRO. Anal. Numér. **14**, 249–277 (1980)
27. Fischer, B., Ramage, A., Silvester, D., Wathen, A.: Minimum residual methods for augmented systems. BIT Numer. Math. **38**, 527–543 (1998)
28. Gatica, G.N., Heuer, N.: Conjugate gradient method for dual-dual mixed formulation. Math. Comput. **71**, 1455–1472 (2001)
29. Girault, V., Raviart, P.: Finite Element Approximation of the Navier–Stokes Equations. Lecture Notes in Math., vol. 749. Springer, New York (1981)
30. Glowinski, R., Pan, T.W., Périaux, J.: Distributed Lagrange multiplier methods for incompressible viscous flow around moving rigid bodies. Comput. Methods Appl. Mech. Eng. **151**, 181–194 (1998)
31. Golub, G., Wathen, A.: An iteration for indefinite systems and its application to the Navier–Stokes equations. SIAM J. Sci. Comput. **19**, 530–539 (1998)
32. Golub, G., Wu, X., Yuan, J.Y.: SOR-like methods for augmented systems. BIT Numer. Math. **41**, 71–85 (2001)
33. Grote, M., Huckle, T.: Parallel preconditioning with sparse approximate inverses. SIAM J. Sci. Comput. **18**, 838–853 (1997)
34. Hu, H.: Direct simulation of flows of solid-liquid mixtures. Int. J. Multiph. Flow **22**, 335–352 (1996)
35. Johnson, A.A., Tezduyar, T.E.: Simulation of multiple spheres falling in a liquid-filled tube. Comput. Methods Appl. Mech. Eng. **134**, 351–373 (1996)
36. Johnson, A.A., Tezduyar, T.E.: 3D simulation of fluid–particle interactions with the number of particles reaching 100. Comput. Methods Appl. Mech. Eng. **145**, 301–321 (1997)
37. Johnson, A.A., Tezduyar, T.E.: Advanced mesh generation and update methods for 3D flow simulations. Comput. Mech. **23**, 130–143 (1999)
38. Johnson, A.A., Tezduyar, T.E.: Methods for 3D computation of fluid-object interactions in spatially-periodic flows. Comput. Methods Appl. Mech. Eng. **190**, 3201–3221 (2001)
39. Keller, C., Gould, N.I.M., Wathen, A.J.: Constraint preconditioning for indefinite linear systems. SIAM J. Matrix Anal. Appl. **21**, 1300–1317 (2000)
40. Klawonn, A.: An optimal preconditioner for a class of saddle point problems with a penalty term. SIAM J. Sci. Comput. **19**, 540–552 (1998)
41. Klawonn, A., Starke, G.: Block triangular preconditioners for nonsymmetric saddle point problems: Field-of-values analysis. Numer. Math. **81**, 577–594 (1999)
42. Knepley, M.: Parallel simulation of the particulate flow problem. Ph.D. thesis, Department of Computer Science, Purdue University (2000)
43. Knepley, M., Sarin, V., Sameh, A.: Parallel simulation of particulate flows. Appeared in Fifth Intl. Symp. on Solving Irregular Structured Problems in Parallel, IRREGULAR 98, LNCS, No. 1457, pp. 226–237, Springer (1998)
44. Krzyzanowski, P.: On block preconditioners for nonsymmetric saddle point problems. SIAM J. Sci. Comput. **23**, 157–169 (2001)
45. Little, L., Saad, Y.: Block LU preconditioners for symmetric and nonsymmetric saddle point problems. Tech. Rep. 1999-104, Minnesota Supercomputer Institute, University of Minnesota (1999)

46. Lou, G.: Some new results for solving linear systems arising from computational fluid dynamics problems. Ph.D. thesis, Department of Computer Science, University of Illinois U-C (1992)
47. Lukšan, L., Vlček, J.: Indefinitely preconditioned inexact Newton method for large sparse equality constrained nonlinear programming problems. Numer. Linear Algebra Appl. **5**, 219–247 (1998)
48. Maury, B.: Characteristics ALE method for the unsteady 3D Navier–Stokes equations with a free surface. Comput. Fluid Dyn. J. **6**, 175–188 (1996)
49. Maury, B.: A many-body lubrication model. C. R. Acad. Sci. Paris **325**, 1053–1058 (1997)
50. Maury, B.: Direct simulations of 2D fluid–particle flows in biperiodic domains. J. Comput. Phys. **156**, 325–351 (1999)
51. Maury, B., Glowinski, R.: Fluid–particle flow: a symmetric formulation. C. R. Acad. Sci. Paris **324**, 1079–1084 (1997)
52. Murphy, M.F., Golub, G.H., Wathen, A.J.: A note on preconditioning for indefinite linear systems. SIAM J. Sci. Comput. **21**, 1969–1972 (2000)
53. Ortega, J.M.: Numerical Analysis: a Second Course. Computer Science and Applied Mathematics Series. Academic Press, San Diego (1972)
54. Perugia, I., Simoncini, V.: An optimal indefinite preconditioner for mixed finite element method. Tech. Rep. 1098, Department of Mathematics, Università de Bologna, Italy (1998)
55. Perugia, I., Simoncini, V.: Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations. Numer. Linear Algebra Appl. **7**, 585–616 (2000)
56. Queck, W.: The convergence factor of preconditioned algorithms of the Arrow-Hurwicz type. SIAM J. Numer. Anal. **26**, 1016–1030 (1989)
57. Rusten, T., Winther, R.: A preconditioned iterative method for saddle point problem. SIAM J. Matrix Anal. Appl. **13**, 887–904 (1992)
58. Saad, Y.: A flexible inner–outer preconditioned GMRES algorithm. SIAM J. Sci. Comput. **14**, 461–469 (1993)
59. Saad, Y., Schultz, M.: GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. **7**, 856–869 (1986)
60. Saad, Y., Suchomel, B.: ARMS: an algebraic recursive multilevel solver for general sparse linear systems. Numer. Linear Algebra Appl. **9**, 359–378 (2002)
61. Sameh, A., Baggag, A.: Parallelism in iterative linear system solvers. In: Proceedings of the Sixth Japan-US Conference on Flow Simulation and Modeling, April 1, 2002
62. Sameh, A., Baggag, A., Wang, X.: Parallel nested iterative schemes for indefinite linear systems. In: Mang, H.A., Rammerstorfer, F.G., Eberhardsteiner, J. (eds.) Proceedings of the Fifth World Congress on Computational Mechanics, (WCCM V). Vienna University of Technology, Austria, July 7–12, 2002. ISBN 3-9501554-0-6
63. Silvester, D., Elman, H., Kay, D., Wathen, A.: Efficient preconditioning of the linearized Navier–Stokes equations for incompressible flow. J. Comput. Appl. Math. **128**, 261–279 (2001)
64. Silvester, D., Wathen, A.: Fast iterative solution of stabilized Stokes systems. part II: Using general block preconditioners. SIAM J. Numer. Anal. **31**, 1352–1367 (1994)
65. Silvester, D., Wathen, A.: Fast and robust solvers for time-discretized incompressible Navier–Stokes equations. Tech. Rep. 27, Department of Mathematics, University of Manchester (1995)
66. Simoncini, V., Szyld, D.: Flexible inner–outer Krylov subspace methods. SIAM J. Numer. Anal. **40**, 2219–2239 (2003)
67. Simoncini, V., Szyld, D.: Theory of inexact Krylov subspace methods and applications to scientific computing. SIAM J. Sci. Comput. **25**, 454–477 (2003)
68. Stewart, G.W.: Introduction to Matrix Computations. Academic Press, San Diego (1973)
69. Tezduyar, T.E.: Stabilized finite element formulations for incompressible flow computations. Adv. Appl. Mech. **28**, 1–44 (1991)
70. Vanderstraeten, D., Knepley, M.: Parallel building blocks for finite element simulations: Application to solid-liquid mixture flows. In: Emerson, D., Ecer, A., Periaux, J., Satofuka, N.

(eds.) Proceedings of Parallel CFD'99 Conf.: Recent Developments and Advances Using Parallel Computers, pp. 133–139. Academic Press, Manchester (1997)
71. Verfürth, R.: A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem. IMA J. Numer. Anal. **4**, 441–455 (1984)
72. Verfürth, R.: A posteriori error estimators for the Stokes equations. Numer. Math. **55**, 309–325 (1989)
73. Wathen, A., Silvester, D.: Fast iterative solution of stabilized Stokes systems. part I: Using simple diagonal preconditioners. SIAM J. Numer. Anal. **30**, 630–649 (1993)
74. Zulehner, W.: A class of smoothers for saddle point problems. Computer **65**, 227–246 (2000)
75. Zulehner, W.: Analysis of iterative methods for saddle point problems: a unified approach. Math. Comput. **71**, 479–505 (2001)