PRECONDITIONING DISCRETE APPROXIMATIONS OF THE REISSNER–MINDLIN PLATE MODEL*

DOUGLAS N. ARNOLD†, RICHARD S. FALK‡, AND RAGNAR WINTHER§

Abstract. We consider iterative methods for the solution of the linear system of equations arising from the mixed finite element discretization of the Reissner–Mindlin plate model. We show how to construct a symmetric positive definite block diagonal preconditioner such that the resulting linear system has spectral condition number independent of both the mesh size h and the plate thickness t. We further discuss how this preconditioner may be implemented and then apply it to efficiently solve this indefinite linear system. Although the mixed formulation of the Reissner–Mindlin problem has a saddle-point structure common to other mixed variational problems, the presence of the small parameter t and the fact that the matrix in the upper left corner of the partition is only positive semidefinite introduces new complications.

Key words. preconditioner, Reissner, Mindlin, plate, finite element

AMS(MOS) subject classifications (1991 revision). 65N30, 65N22, 65F10, 73V05

1. Introduction. The Reissner–Mindlin plate model can be formulated as a saddle point problem and discretized by mixed finite element methods. The resulting linear algebraic system is symmetric and nonsingular, but indefinite. In this paper we show how this system can be efficiently solved by preconditioned iterative methods. In particular, we will establish bounds on the number of iterations necessary to achieve any desired error reduction factor (in an appropriate norm) with the bounds independent of both the discretization parameter h and the plate thickness t.

In order to understand our approach, consider a problem

$$\mathcal{A}x = f \tag{1.1}$$

arising from the discretization of a well-posed linear boundary value problem by a stable discretization scheme. We assume that $\mathcal{A}$ is self-adjoint, but not necessarily positive definite, and defines an isomorphism from an appropriate Banach space X to its dual X^*. Note that the operator $\mathcal{A}$ and the space X depend on the discretization parameters and perhaps on other parameters as well (for example, on the thickness t in the Reissner–Mindlin model), but we suppose that we have bounds on $\|\mathcal{A}\|_{\mathcal{L}(X,X^*)}$ and $\|\mathcal{A}^{-1}\|_{\mathcal{L}(X^*,X)}$ which are independent of these parameters.

*The first author was supported NSF grants DMS-9205300 and DMS-9500672 and by the Institute for Mathematics and its Applications. The second author was supported by NSF grant DMS-9403552. The third author was supported by The Norwegian Research Council under grants 100331/431 and STP.29643.

†Department of Mathematics, Penn State, University Park, PA 16802. dna@math.psu.edu

‡Department of Mathematics, Rutgers University, New Brunswick, NJ 08903. falk@math.rutgers.edu

§Department of Informatics, University of Oslo, Oslo, Norway. ragnar@ifi.uio.no

In order to solve (1.1), we will use the minimum residual method or another iterative scheme with similar properties, preconditioning with a positive definite self-adjoint operator $\mathcal{B} : X^* \to X$. Such an iterative scheme is efficient if the action of the operator $\mathcal{B}$ may be computed efficiently and if the magnitude of the eigenvalues of $\mathcal{B}\mathcal{A}$ can be bounded above and below by positive constants (cf. § 5 below). It is easy to see that such spectral bounds follow directly from the bounds on $\|\mathcal{A}\|_{\mathcal{L}(X,X^*)}$ and $\|\mathcal{A}^{-1}\|_{\mathcal{L}(X^*,X)}$ and on bounds for $\|\mathcal{B}\|_{\mathcal{L}(X^*,X)}$ and $\|\mathcal{B}^{-1}\|_{\mathcal{L}(X,X^*)}$. Thus to efficiently solve (1.1), we simply require a computable positive definite operator $\mathcal{B}$ for which we can bound the norm and the norm of its inverse uniformly in the relevant parameters. We remark that the preconditioner $\mathcal{B}$ can be constructed without reference to the detailed structure of the operator $\mathcal{A}$, but depends only on the Banach space X.

In case $\mathcal{A}$ is associated with a differential system, as in the Reissner–Mindlin model, the space X will be a Cartesian product $X_1 \times \cdots X_n$. Therefore it is easy to construct $\mathcal{B}$ if computable self-adjoint positive definite operators $\mathcal{B}_i : X_i^* \to X_i$ are available; we simply set

$$\mathcal{B} = \begin{pmatrix} \mathcal{B}_1 & & \\ & \ddots & \\ & & \mathcal{B}_n \end{pmatrix}.$$

The $\mathcal{B}_i$ will typically be preconditioners for simpler subproblems.

Like the biharmonic plate model, the Reissner–Mindlin model is a two-dimensional plate model which approximates the behavior of a thin linearly elastic three-dimensional body using unknowns and equations defined only on the middle surface, Ω, of the plate. The basic variables of the model are the transverse displacement ω and the rotation vector $\boldsymbol{\phi}$ which solve the system of partial differential equations

$$-\,\mathbf{div}\, \mathcal{C}\, \mathcal{E}\, \boldsymbol{\phi} + \lambda t^{-2}(\boldsymbol{\phi} - \mathbf{grad}\,\omega) = 0, \tag{1.2}$$

$$\lambda t^{-2}(-\,\Delta\,\omega + \operatorname{div} \boldsymbol{\phi}) = g, \tag{1.3}$$

on Ω together with suitable boundary conditions. For the hard clamped plate, which we consider throughout this paper, these are $\omega = 0$, $\boldsymbol{\phi} = 0$. In (1.2)–(1.3), g is the scaled transverse loading function, t is the plate thickness, $\mathcal{E}\,\boldsymbol{\phi}$ is the symmetric part of the gradient of $\boldsymbol{\phi}$, and the scalar constant λ and constant tensor $\mathcal{C}$ depend on the material properties of the body. Precise definitions are given in the next section. A variational formulation of this system states that the solution $(\boldsymbol{\phi}, \omega)$ minimizes the total energy of the plate, which is given by

$$E(\boldsymbol{\phi}, \omega) = \frac{1}{2} \int_\Omega (\mathcal{C}\, \mathcal{E}\, \boldsymbol{\phi}) : (\mathcal{E}\, \boldsymbol{\phi})\, dx + \frac{\lambda t^{-2}}{2} \int_\Omega |\boldsymbol{\phi} - \mathbf{grad}\,\omega|^2\, dx - \int_\Omega g\omega\, dx$$

over $\mathring{\boldsymbol{H}}^1(\Omega) \times \mathring{H}^1(\Omega)$.

An advantage of the Reissner–Mindlin model over the biharmonic plate model is that the energy involves only first derivatives of the unknowns and so conforming finite element approximations require the use of merely continuous finite element spaces rather than the C^1 spaces required for the biharmonic model. However, for many choices of finite element spaces, severe difficulties arise due to the presence of the small parameter t. If the finite element subspaces are not properly related, the phenomenon of "locking" occurs, causing a deterioration in the approximation as the plate thickness t approaches zero. A key step in understanding and overcoming locking is passage to a mixed formulation of the Reissner–Mindlin model. The mixed formulation may be derived from the alternative system of differential equations

$$-\operatorname{\mathbf{div}} \mathcal{C}\, \mathcal{E}\, \boldsymbol{\phi} - \boldsymbol{\zeta} = 0, \tag{1.4}$$

$$-\operatorname{div} \boldsymbol{\zeta} = g, \tag{1.5}$$

$$-\boldsymbol{\phi} + \operatorname{\mathbf{grad}} \omega - \lambda^{-1} t^2 \boldsymbol{\zeta} = 0, \tag{1.6}$$

arising from (1.2)–(1.3) through the introduction of the shear stress $\boldsymbol{\zeta} = \lambda t^{-2}(\operatorname{\mathbf{grad}} \omega - \boldsymbol{\phi})$. A variational statement is that $(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta}) \in \mathring{\boldsymbol{H}}^1(\Omega) \times \mathring{H}^1(\Omega) \times \boldsymbol{L}^2(\Omega)$ is a critical point of the mixed Lagrangian

$$L(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta}) = \frac{1}{2} \int_\Omega (\mathcal{C}\, \mathcal{E}\, \boldsymbol{\phi}) : (\mathcal{E}\, \boldsymbol{\phi})\, dx - \int_\Omega \boldsymbol{\zeta} \cdot (\boldsymbol{\phi} - \operatorname{\mathbf{grad}} \omega)\, dx - \frac{\lambda^{-1} t^2}{2} \int_\Omega |\boldsymbol{\zeta}|^2\, dx - \int_\Omega g\omega\, dx.$$

This is a saddle-point principle, and consequently the linear equations arising from its discretization are indefinite. Many finite element methods which have been derived and analyzed for this mixed formulation can be implemented in terms of the primitive variables $\boldsymbol{\phi}$ and ω only (that is $\boldsymbol{\zeta}$ can be eliminated at the discrete level) and therefore lead to a positive definite linear system. However, we do not know how to derive efficient preconditioners for the solution of these systems (uniformly in t). In this paper, we shall propose preconditioned iterative methods for the full indefinite mixed system.

The application of multigrid methods for the solution of discrete Reissner–Mindlin systems and related problems has been considered by several authors. Braess and Blömer [3] consider the analogous problem in one-space dimension, the Timoshenko beam model, and show that if multigrid methods are applied directly to the discrete positive definite system corresponding to analogues of (1.2), (1.3), then the convergence rate deteriorates as the beam thickness tends to zero. By contrast, they formulate a multigrid W-cycle algorithm for the mixed system, using a smoother based on the normal equations, and show that the convergence rate is independent of the beam thickness t and discretization parameter h. Peisker [17] formulates a multigrid method for a family of discretizations of the system corresponding to (1.2), (1.3) using polynomials of degree 2 or greater. She shows that if t is no less than h, and if sufficiently many smoothing steps are made, then the method converges with a rate independent of t and h. She also discusses the application to problems with t smaller than h using an additional iteration. Huang [15] studies a different mixed formulation of the problem as a perturbed Stokes-like system which arises from the Helmholtz decomposition of $\boldsymbol{\zeta}$, and uses this formulation to devise and analyze a multigrid algorithm for the method proposed in [1]. More recently, a general framework which includes the methods of [3] and [15] is given in [8].

The approach of this paper is different from that taken in the papers described above, since our strategy is to use symmetric positive definite block diagonal preconditioners for the mixed system, where the blocks correspond to preconditioners for simpler subproblems. These simpler preconditioners may be constructed, for example, using multigrid or domain decomposition techniques.

Other authors have considered the use of symmetric positive definite block diagonal preconditioners for the indefinite algebraic systems arising from certain saddle point problems, such as the mixed formulation of scalar second order elliptic equations and the Stokes equations. See Bramble and Pasciak [5], Klawonn [16], Rusten and Winther [19], [20], Silvester and Wathen [21], and Wathen and Silvester [24]. In designing and analyzing their preconditioners, these authors have exploited the fact that for these problems the upper left hand corner of the coefficient matrix, A_h, is positive definite, and so their techniques don't apply directly to the equations arising from the Reissner–Mindlin system. Other approaches to the design of iterative methods for algebraic systems arising from saddle point problems include the the inexact Uzawa algorithms analyzed by Elman and Golub [13] and Bramble, Pasciak, and Vassilev [6], and a method based on a positive definite reformulation is discussed by Bramble and Pasciak [4]. Again, these analyses rely on the definiteness of the upper left hand corner of the coefficient matrix and so would have to be modified for use with Reissner–Mindlin.

An outline of the paper is as follows. In the next section, we collect some preliminary results about various formulations of the boundary value problem for the Reissner–Mindlin model. Appropriate Hilbert spaces are defined for the data and solution and an isomorphism theorem is stated relating the two. This result uses t-dependent norms and is, as far as we know, new. In § 3 we discuss finite element methods for the Reissner–Mindlin model. We pose three hypotheses which delimit precisely the class of methods to which our results apply, and show that methods in the literature which have been proven to be free of locking satisfy these hypotheses. In § 4, we state and prove the isomorphism theorem for the class of finite element schemes under consideration. Using this result, we then discuss in §§ 5–7 how to precondition the linear systems resulting from these approximation schemes. Finally, in § 8, we report the results of numerical experiments using some of the preconditioning methods developed in this paper.

2. Preliminaries. We begin this section by recalling the sum and intersection construction for Hilbert spaces. If Hilbert spaces X and Y are both continuously included in some larger Hilbert space, then the intersection $X \cap Y$ and the sum $X + Y$ are themselves

Hilbert spaces with the norms

$$\|z\|_{X\cap Y} = (\|z\|_X^2 + \|z\|_Y^2)^{1/2} \text{ and } \|z\|_{X+Y} = \inf_{\substack{x\in X, y\in Y\\ x+y=z}} (\|x\|_X^2 + \|y\|_Y^2)^{1/2}.$$

If in addition, $X\cap Y$ is dense in both X and Y, then the dual spaces X^* and Y^* may be viewed as subspaces of $(X\cap Y)^*$. Moreover we have the following result (see [2, § 2.7] for a proof in the Banach space context).

THEOREM 2.1. *The dual space* $(X\cap Y)^* = X^* + Y^*$ *and the dual norm coincides with the sum norm:*

$$\sup_{z\in X\cap Y} \frac{\langle z^*, z\rangle}{\|z\|_{X\cap Y}} = \|z^*\|_{X^*+Y^*}, \quad \textit{for all } z^* \in (X\cap Y)^*.$$

We next define the notation to be used. For simplicity, we assume that Ω is a polygonal domain in $\mathbb{R}^2$. Since the domain Ω is fixed throughout the paper, we shall adopt notation which omits explicit reference to Ω. Hence, we will use H^m to denote the Sobolev space of functions with m derivatives in $L^2(\Omega)$, while $\mathring{H}^m$ denotes the subspace obtained as the closure in H^m of $C_0^\infty(\Omega)$. The dual space of $\mathring{H}^m$ with respect to the L^2 inner product will be denoted by H^{-m}. A space written in boldface denotes the 2-vector-valued analogue of the corresponding scalar-valued space. For both the scalar- and the vector-valued Sobolev space of order m, we use $\|\cdot\|_m$ to denote the norm. The notation $(\,\cdot\,,\,\cdot\,)$ is used to denote the L^2 inner product of scalar, vector, and matrix valued functions.

We now present the weak formulations of the systems (1.2)–(1.3) and (1.4)–(1.6). In these equations $t > 0$ is the plate thickness and the positive constant $\lambda = Ek/2(1+\nu)$ with E the Young's modulus, ν the Poisson ratio, and k the shear correction factor. For all 2×2 symmetric matrices $\boldsymbol{\tau}$, $\mathcal{C}\boldsymbol{\tau}$ is the 2×2 symmetric matrix defined by

$$\mathcal{C}\boldsymbol{\tau} = \frac{E}{12(1-\nu^2)}[(1-\nu)\boldsymbol{\tau} + \nu\,\mathrm{tr}(\boldsymbol{\tau})I],$$

where $\mathrm{tr}(\boldsymbol{\tau})$ denotes the trace of $\boldsymbol{\tau}$, so $\mathcal{C}$ defines a symmetric positive definite operator from the space of 2×2 symmetric matrices to itself. The function g, which represents the scaled load function, is assumed to belong to H^{-1}. In the interest of notational simplicity and without loss of generality, we shall set $\lambda = 1$ from this point on. (All the results which follow may be translated back to the original case by replacing t with $\lambda^{-1/2}t$.)

The weak formulation of the system (1.2)–(1.3) is:

Problem (P): Find $\boldsymbol{\phi} \in \mathring{\boldsymbol{H}}^1$, $\omega \in \mathring{H}^1$, satisfying

$$(\mathcal{C}\,\mathcal{E}\,\boldsymbol{\phi}, \mathcal{E}\,\boldsymbol{\psi}) + t^{-2}(\boldsymbol{\phi} - \mathbf{grad}\,\omega, \boldsymbol{\psi} - \mathbf{grad}\,\mu) = (g,\mu) \quad \text{for all } \boldsymbol{\psi} \in \mathring{\boldsymbol{H}}^1, \mu \in \mathring{H}^1.$$

The existence of a unique solution to this problem is straightforward. Introducing the shear stress $\boldsymbol{\zeta}$ as in (1.4)–(1.6), we obtain the mixed weak formulation:

Problem (M): Find $\boldsymbol{\phi} \in \mathring{\boldsymbol{H}}^1$, $\omega \in \mathring{H}^1$, $\boldsymbol{\zeta} \in \boldsymbol{L}^2$ satisfying

$$(\mathcal{C}\,\mathcal{E}\,\boldsymbol{\phi}, \mathcal{E}\,\boldsymbol{\psi}) - (\boldsymbol{\zeta}, \boldsymbol{\psi} - \mathbf{grad}\,\mu) = (g, \mu) \quad \text{for all } \boldsymbol{\psi} \in \mathring{\boldsymbol{H}}^1, \mu \in \mathring{H}^1,$$
$$-(\boldsymbol{\phi} - \mathbf{grad}\,\omega, \boldsymbol{\eta}) - t^2(\boldsymbol{\zeta}, \boldsymbol{\eta}) = 0 \quad \text{for all } \boldsymbol{\eta} \in \boldsymbol{L}^2.$$

A proof of existence, uniqueness, and a priori estimates for this system based on Brezzi's theory of saddle-point problems may be found in [9].

We note that Problem (M), unlike Problem (P), has a sense when $t = 0$. Indeed, for $t = 0$ one easily obtains that $\boldsymbol{\phi} = \mathbf{grad}\,\omega$ and $\boldsymbol{\zeta} = E[12(1-\nu^2)]^{-1}\,\mathbf{grad}\,\Delta\omega$, where $\omega \in \mathring{H}^2$ satisfies

$$(\mathcal{C}\,\mathcal{E}\,\mathbf{grad}\,\omega, \mathcal{E}\,\mathbf{grad}\,\mu) = -(g, \mu) \quad \text{for all } \mu \in \mathring{H}^2.$$

This is a weak formulation of the biharmonic model for the clamped plate. However, at this limit the regularity $\boldsymbol{\zeta} \in \boldsymbol{L}^2$ is lost, and the proper statement of Problem (M) places $\boldsymbol{\zeta}$ and $\boldsymbol{\eta}$ in the space $\boldsymbol{H}^{-1}(\mathrm{div})$. This space is defined as the dual of

$$\mathring{\boldsymbol{H}}(\mathrm{rot}) = \{\boldsymbol{\eta} \in \boldsymbol{L}^2 : \mathrm{rot}\,\boldsymbol{\eta} \in L^2, \boldsymbol{\eta} \cdot \boldsymbol{s} = 0 \text{ on } \partial\Omega\},$$

where the norm in $\mathring{\boldsymbol{H}}(\mathrm{rot})$ is given by

$$\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}(\mathrm{rot})} = \left(\|\boldsymbol{\eta}\|_0^2 + \|\,\mathrm{rot}\,\boldsymbol{\eta}\|_0^2\right)^{1/2}.$$

Here $\boldsymbol{s}$ is the unit tangent to $\partial\Omega$ and $\mathrm{rot}\,\boldsymbol{\eta} = \partial\boldsymbol{\eta}_1/\partial y - \partial\boldsymbol{\eta}_2/\partial x$. It can be shown that the dual space

$$\boldsymbol{H}^{-1}(\mathrm{div}) = \{\boldsymbol{\eta} \in \boldsymbol{H}^{-1} : \mathrm{div}\,\boldsymbol{\eta} \in H^{-1}\},$$

and that the norm

$$\boldsymbol{\eta} \mapsto \left(\|\boldsymbol{\eta}\|_{-1}^2 + \|\,\mathrm{div}\,\boldsymbol{\eta}\|_{-1}^2\right)^{1/2}$$

is equivalent to the dual norm

$$\|\boldsymbol{\zeta}\|_{\boldsymbol{H}^{-1}(\mathrm{div})} = \sup_{\boldsymbol{\eta} \in \mathring{\boldsymbol{H}}(\mathrm{rot})} \frac{(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}(\mathrm{rot})}}.$$

In order to state the mixed formulation in a manner which is valid up to and including the limit $t = 0$ and to describe the regularity of solutions, we define some Hilbert spaces with norms depending on t. For $t > 0$ the space $t \cdot \boldsymbol{L}^2$ is simply the space $\boldsymbol{L}^2$ except with norm multiplied by t. The space $\boldsymbol{H}^{-1}(\mathrm{div}) \cap t \cdot \boldsymbol{L}^2$ is then defined as an intersection space. As a set it coincides with $\boldsymbol{L}^2$, but has norm given by

$$\|\boldsymbol{\eta}\|_{\boldsymbol{H}^{-1}(\mathrm{div}) \cap t \cdot \boldsymbol{L}^2} = \left(\|\boldsymbol{\eta}\|_{\boldsymbol{H}^{-1}(\mathrm{div})}^2 + t^2\|\boldsymbol{\eta}\|_0^2\right)^{1/2}.$$

In view of Theorem 2.1, its dual space is $\mathring{\boldsymbol{H}}(\mathrm{rot}) + t^{-1} \cdot \boldsymbol{L}^2$, which again coincides with $\boldsymbol{L}^2$ as a set, but has norm

$$\|\boldsymbol{\zeta}\|_{\mathring{\boldsymbol{H}}(\mathrm{rot})+t^{-1}\cdot\boldsymbol{L}^2} = \sup_{\boldsymbol{\eta}\in\boldsymbol{L}^2} \frac{(\boldsymbol{\zeta},\boldsymbol{\eta})}{\|\boldsymbol{\eta}\|_{\boldsymbol{H}^{-1}(\mathrm{div})\cap t\cdot\boldsymbol{L}^2}} = \inf_{\boldsymbol{\zeta}_1+\boldsymbol{\zeta}_2=\boldsymbol{\zeta}} \left(\|\boldsymbol{\zeta}_1\|^2_{\mathring{\boldsymbol{H}}(\mathrm{rot})} + t^{-2}\|\boldsymbol{\zeta}_2\|_0^2\right)^{1/2}.$$

When $t = 0$, these spaces become $\boldsymbol{H}^{-1}(\mathrm{div})$ and $\mathring{\boldsymbol{H}}(\mathrm{rot})$, respectively. Hence, if we replace $\boldsymbol{L}^2$ by $\boldsymbol{H}^{-1}(\mathrm{div}) \cap t \cdot \boldsymbol{L}^2$ in Problem (M), the formulation is valid for $t \geq 0$. In fact we shall consider a slightly generalized system:

Problem (G): Find $\boldsymbol{\phi} \in \mathring{\boldsymbol{H}}^1$, $\omega \in \mathring{H}^1$, $\boldsymbol{\zeta} \in \boldsymbol{H}^{-1}(\mathrm{div}) \cap t \cdot \boldsymbol{L}^2$ satisfying

$$\begin{aligned}
(\mathcal{C}\,\mathcal{E}\,\boldsymbol{\phi}, \mathcal{E}\,\boldsymbol{\psi}) - (\boldsymbol{\zeta}, \boldsymbol{\psi}) &= (\boldsymbol{f}, \boldsymbol{\psi}), \quad \text{for all } \boldsymbol{\psi} \in \mathring{\boldsymbol{H}}^1, \\
(\boldsymbol{\zeta}, \mathbf{grad}\,\mu) &= (g, \mu), \quad \text{for all } \mu \in \mathring{H}^1, \\
-(\boldsymbol{\phi} - \mathbf{grad}\,\omega, \boldsymbol{\eta}) - t^2(\boldsymbol{\zeta}, \boldsymbol{\eta}) &= (\boldsymbol{j}, \boldsymbol{\eta}), \quad \text{for all } \boldsymbol{\eta} \in \boldsymbol{H}^{-1}(\mathrm{div}) \cap t \cdot \boldsymbol{L}^2,
\end{aligned}$$

where $\boldsymbol{f} \in \boldsymbol{H}^{-1}$ and $\boldsymbol{j} \in \mathring{\boldsymbol{H}}(\mathrm{rot}) + t^{-1} \cdot \boldsymbol{L}^2$.

Let X_t denote the product space

$$X_t = \mathring{\boldsymbol{H}}^1 \times \mathring{H}^1 \times \boldsymbol{H}^{-1}(\mathrm{div}) \cap t \cdot \boldsymbol{L}^2.$$

For $t > 0$, X_t coincides with $\mathring{\boldsymbol{H}}^1 \times \mathring{H}^1 \times \boldsymbol{L}^2$ as a set, but its norm,

$$\|(\boldsymbol{\psi}, \mu, \boldsymbol{\eta})\|_{X_t} = (\|\boldsymbol{\psi}\|_1^2 + \|\mu\|_1^2 + \|\boldsymbol{\eta}\|^2_{\boldsymbol{H}^{-1}(\mathrm{div})\cap t\cdot\boldsymbol{L}^2})^{1/2}$$

is t-dependent. Let X_t^* denote the dual space, $\boldsymbol{H}^{-1} \times H^{-1} \times \mathring{\boldsymbol{H}}(\mathrm{rot}) + t^{-1} \cdot \boldsymbol{L}^2$ with the dual norm:

$$\|(\boldsymbol{\psi}, \mu, \boldsymbol{\eta})\|_{X_t^*} = (\|\boldsymbol{\psi}\|_{-1}^2 + \|\mu\|_{-1}^2 + \|\boldsymbol{\eta}\|^2_{\mathring{\boldsymbol{H}}(\mathrm{rot})+t^{-1}\cdot\boldsymbol{L}^2})^{1/2}.$$

The following isomorphism theorem holds for Problem (G).

THEOREM 2.2. *Let $(\boldsymbol{f}, g, \boldsymbol{j}) \in X_t^*$ be given. Then there is a unique solution $(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta}) \in X_t$ to Problem (G). Moreover there exist positive constants c_1 and c_2 independent of t, $\boldsymbol{f}$, g, and $\boldsymbol{j}$ such that for $0 \leq t \leq 1$,*

$$c_1\|(\boldsymbol{f}, g, \boldsymbol{j})\|_{X_t^*} \leq \|(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta})\|_{X_t} \leq c_2\|(\boldsymbol{f}, g, \boldsymbol{j})\|_{X_t^*}.$$

Since we shall prove the discrete analogue of this theorem, which is somewhat more complicated but follows the same framework, in § 5, we omit the proof here. For the remainder of the paper we shall assume, as in the hypothesis above, that $0 \leq t \leq 1$.

Before turning to the description of the discretization schemes, we motivate our construction of preconditioners by discussing the preconditioning of the continuous Problem (G). Let $\mathcal{A}_t : X_t \to X_t^*$ be the continuous operator so that Problem (G) can be written in the form

$$\mathcal{A}_t \begin{pmatrix} \boldsymbol{\phi} \\ \omega \\ \boldsymbol{\zeta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ g \\ \boldsymbol{j} \end{pmatrix}$$

and Theorem 2.2 provides bounds on $\|\mathcal{A}_t\|_{\mathcal{L}(X_t, X_t^*)}$ and $\|\mathcal{A}_t^{-1}\|_{\mathcal{L}(X_t^*, X_t)}$ uniform in t. Note that $\mathcal{A}_t$ is given by the differential operator

$$\mathcal{A}_t = \begin{pmatrix} -\operatorname{\mathbf{div}} \mathcal{C}\,\mathcal{E} & 0 & -\boldsymbol{I} \\ 0 & 0 & -\operatorname{div} \\ -\boldsymbol{I} & \operatorname{\mathbf{grad}} & -t^2 \boldsymbol{I} \end{pmatrix}.$$

As discussed in the introduction, we require a self-adjoint positive definite operator $\mathcal{B}_t : X_t^* \to X_t$ for which $\|\mathcal{B}_t\|_{\mathcal{L}(X_t^*, X_t)}$ and $\|\mathcal{B}_t^{-1}\|_{\mathcal{L}(X_t, X_t^*)}$ are bounded independently of t. If $\boldsymbol{L} : \boldsymbol{H}^{-1} \to \mathring{\boldsymbol{H}}^1$, $M : H^{-1} \to \mathring{H}^1$, and $\boldsymbol{N}_t : \mathring{\boldsymbol{H}}(\mathrm{rot}) + t^{-1} \cdot \boldsymbol{L}^2 \to \boldsymbol{H}^{-1}(\mathrm{div}) \cap t \cdot \boldsymbol{L}^2$ are self-adjoint positive definite operators for which analogous bounds hold, then the desired preconditioner can be taken to be

$$\mathcal{B}_t = \begin{pmatrix} \boldsymbol{L} & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & \boldsymbol{N}_t \end{pmatrix}.$$

At the discrete level, the operators $\boldsymbol{L}_h$ and M_h can simply be taken to be any standard preconditioners for second order elliptic operators. Note that the operator $\boldsymbol{N}_t$, in contrast to the operators $\boldsymbol{L}$ and M, decreases regularity. In particular, for $t = 0$, $\boldsymbol{N}_0$ is required to be an isomorphism from $\mathring{\boldsymbol{H}}(\mathrm{rot})$ into $\boldsymbol{H}^{-1}(\mathrm{div})$. The natural choice for such an operator is given by $\boldsymbol{I} + \operatorname{\mathbf{curl}} \operatorname{rot}$, where the curl operator is defined by

$$\operatorname{\mathbf{curl}} = \begin{pmatrix} -\partial/\partial y \\ \partial/\partial x \end{pmatrix}.$$

At the discrete level we will want $\boldsymbol{N}_{t,h}$ to be an approximation of this differential operator (at least for $t = 0$).

With this motivation, in the next section we describe the discretization scheme and establish a discrete analogue of the isomorphism theorem.

3. Finite element approximation schemes. We now turn to the description of approximation schemes. For $t > 0$ most finite element approximations which have been proposed for the Reissner–Mindlin system can be expressed in the form:

Problem (P_h): Find $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$, $\omega_h \in W_h$, satisfying

$$(\mathcal{C}\,\mathcal{E}\,\boldsymbol{\phi}_h, \mathcal{E}\,\boldsymbol{\psi}) + t^{-2}(\boldsymbol{R}_h \boldsymbol{\phi}_h - \operatorname{\mathbf{grad}}_h \omega_h, \boldsymbol{R}_h \boldsymbol{\psi} - \operatorname{\mathbf{grad}}_h \mu) = (g, \mu), \quad \text{for all } \boldsymbol{\psi} \in \boldsymbol{V}_h, \mu \in W_h.$$

Here $\boldsymbol{V}_h \subset \mathring{\boldsymbol{H}}^1$ and $W_h \subset L^2$ is a finite element approximation of $\mathring{H}^1$ which may or may not be conforming. The reduction operator $\boldsymbol{R}_h$ maps $\boldsymbol{V}_h$ into a third finite element space $\boldsymbol{\Gamma}_h \subset \boldsymbol{L}^2$ and $\mathbf{grad}_h : W_h \to \boldsymbol{\Gamma}_h$ denotes a discrete gradient operator which we assume to be injective on W_h (in most cases W_h is a conforming approximation of $\mathring{H}^1$ and $\mathbf{grad}_h$ is the ordinary gradient). The operator $\boldsymbol{R}_h$ and the additional space $\boldsymbol{\Gamma}_h$ are introduced in order to avoid the locking problem mentioned in the introduction. Set $X_{t,h} = \boldsymbol{V}_h \times W_h \times \boldsymbol{\Gamma}_h$. As a set, this space is independent of t, but a t-dependent norm will be defined below. Introducing $\boldsymbol{\zeta}_h = t^{-2}(\mathbf{grad}_h\, \omega_h - \boldsymbol{R}_h \boldsymbol{\phi}_h)$ and generalizing this problem to allow more general data $(\boldsymbol{f}_h, g_h, \boldsymbol{j}_h) \in X_{t,h}$, we obtain the following discrete problem.

Problem (G_h): Find $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$, $\omega_h \in W_h$, $\boldsymbol{\zeta}_h \in \boldsymbol{\Gamma}_h$ satisfying

$$(\mathcal{C}\, \mathcal{E}\, \boldsymbol{\phi}_h, \mathcal{E}\, \boldsymbol{\psi}) - (\boldsymbol{\zeta}_h, \boldsymbol{R}_h \boldsymbol{\psi}) = (\boldsymbol{f}_h, \boldsymbol{\psi}), \quad \text{for all } \boldsymbol{\psi} \in \boldsymbol{V}_h, \tag{3.1}$$

$$(\boldsymbol{\zeta}_h, \mathbf{grad}_h\, \nu) = (g_h, \nu), \quad \text{for all } \nu \in W_h, \tag{3.2}$$

$$-(\boldsymbol{R}_h \boldsymbol{\phi}_h - \mathbf{grad}_h\, \omega_h, \boldsymbol{\eta}) - t^2(\boldsymbol{\zeta}_h, \boldsymbol{\eta}) = (\boldsymbol{j}_h, \boldsymbol{\eta}), \quad \text{for all } \boldsymbol{\eta} \in \boldsymbol{\Gamma}_h. \tag{3.3}$$

This problem can be expressed compactly as

$$\mathcal{A}_{t,h} \begin{pmatrix} \boldsymbol{\phi}_h \\ \omega_h \\ \boldsymbol{\zeta}_h \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_h \\ g_h \\ \boldsymbol{j}_h \end{pmatrix}, \tag{3.4}$$

where the operator $\mathcal{A}_{t,h} : X_h \to X_h$ is defined implicitly by the left hand side of problem (G_h). Specifically the operator $\mathcal{A}_{t,h}$ has the form

$$\mathcal{A}_{t,h} = \begin{pmatrix} A_h & B_h^* \\ B_h & -t^2 I \end{pmatrix},$$

where $A_h : \boldsymbol{V}_h \times W_h \to \boldsymbol{V}_h \times W_h$ and $B_h : \boldsymbol{V}_h \times W_h \to \boldsymbol{\Gamma}_h$ are given by

$$\left(A_h \begin{pmatrix} \boldsymbol{\phi} \\ \omega \end{pmatrix}, \begin{pmatrix} \boldsymbol{\psi} \\ \mu \end{pmatrix} \right) = (\mathcal{C}\, \mathcal{E}\, \boldsymbol{\phi}, \mathcal{E}\, \boldsymbol{\psi}) \quad \text{for all } (\boldsymbol{\phi}, \omega), (\boldsymbol{\psi}, \mu) \in \boldsymbol{V}_h \times W_h,$$

$$\left(B_h \begin{pmatrix} \boldsymbol{\phi} \\ \omega \end{pmatrix}, \boldsymbol{\eta} \right) = -(R_h \boldsymbol{\phi}, \boldsymbol{\eta}) + (\mathbf{grad}_h\, \omega, \boldsymbol{\eta}) \quad \text{for all } (\boldsymbol{\phi}, \omega) \in \boldsymbol{V}_h \times W_h, \quad \boldsymbol{\eta} \in \boldsymbol{\Gamma}_h,$$

and B_h^* is the L^2-adjoint of B_h.

The operator A_h is symmetric and positive semidefinite with respect to the L^2 inner product. In fact, Korn's inequality implies that $\boldsymbol{\phi} \mapsto (\mathcal{C}\, \mathcal{E}\, \boldsymbol{\phi}, \mathcal{E}\, \boldsymbol{\phi})^{1/2}$ is a norm on $\mathring{\boldsymbol{H}}^1$ equivalent to the usual one. However, A_h is only positive semidefinite since its kernel $\{0\} \times W_h$ is nonzero. The operator $\mathcal{A}_{t,h}$ is self adjoint with respect to the L^2 inner product on X_h. Furthermore, under appropriate assumptions on the finite element spaces, which are introduced in the next section, $\mathcal{A}_{t,h}$ is nonsingular.

The proof of the discrete version of Theorem 2.2 which we shall give in § 5 will require that the discretization scheme satisfy certain abstract hypotheses. These are mostly the same hypotheses that are needed to prove that the scheme is free of locking. After stating them, we shall give several examples of finite element spaces which satisfy them.

The mixed finite element method given by Problem (M_h) is determined by the specification of the finite element spaces $\boldsymbol{V}_h \subset \mathring{\boldsymbol{H}}^1$, $W_h \subset L^2$, and $\boldsymbol{\Gamma}_h \subset \boldsymbol{L}^2$, and the operators $\mathbf{grad}_h : W_h \to \boldsymbol{\Gamma}_h$ (assumed to be injective on W_h) and $\boldsymbol{R}_h : \boldsymbol{V}_h \to \boldsymbol{\Gamma}_h$. In addition, we require, for the analysis only, a space $Q_h \subset L_0^2$ (the subspace of L^2 consisting of functions with mean value zero) and an operator $\mathbf{curl}_h : Q_h \to \boldsymbol{\Gamma}_h$. Adjoint operators $\operatorname{div}_h : \boldsymbol{\Gamma}_h \to W_h$ and $\operatorname{rot}_h : \boldsymbol{\Gamma}_h \to Q_h$ are defined by

$$(\operatorname{div}_h \boldsymbol{\zeta}, \mu) = -(\boldsymbol{\zeta}, \mathbf{grad}_h \mu) \quad \text{for all } \boldsymbol{\zeta} \in \boldsymbol{\Gamma}_h, \mu \in W_h,$$

and $\operatorname{rot}_h : \boldsymbol{\Gamma}_h \to Q_h$ by

$$(\operatorname{rot}_h \boldsymbol{\zeta}, q) = (\boldsymbol{\zeta}, \mathbf{curl}_h q) \quad \text{for all } \boldsymbol{\zeta} \in \boldsymbol{\Gamma}_h, q \in Q_h. \tag{3.5}$$

We then make the following hypotheses:

(H1) (Discrete Helmholtz decomposition)

$$\boldsymbol{\Gamma}_h = \mathbf{grad}_h W_h \oplus \mathbf{curl}_h Q_h,$$

the decomposition being orthogonal with respect to the L^2 inner product.

(H2) There exists a constant C_1 independent of h such that

$$\|\boldsymbol{R}_h \boldsymbol{\psi}\|_0 + \|\operatorname{rot}_h \boldsymbol{R}_h \boldsymbol{\psi}\|_0 \le C_1 \|\boldsymbol{\psi}\|_1 \quad \text{for all } \boldsymbol{\psi} \in \boldsymbol{V}_h.$$

(H3) There exists a positive constant C_2 independent of h such that

$$\sup_{\boldsymbol{\psi} \in \boldsymbol{V}_h} \frac{(\mathbf{curl}_h p, \boldsymbol{R}_h \boldsymbol{\psi})}{\|\boldsymbol{\psi}\|_1} \ge C_2^{-1} \|p\|_0 \quad \text{for all } p \in Q_h.$$

Note that by (3.5) and (H1), $\operatorname{rot}_h \mathbf{grad}_h r = 0$ for all $r \in W_h$. A useful estimate, which follows directly from (H2) and (H3) is that

$$\|p\|_0 \le C \|\mathbf{curl}_h p\|_0 \quad \text{for all } p \in Q_h, \tag{3.6}$$

where C is a constant independent of h.

The statement of the discrete isomorphism theorem will involve the use of several mesh dependent norms based on the discrete operators $\mathbf{grad}_h$, $\mathbf{curl}_h$, div_h, and rot_h. We define for $\boldsymbol{\phi} \in \boldsymbol{V}_h$, $\omega \in W_h$, and $\boldsymbol{\zeta} \in \boldsymbol{\Gamma}_h$,

$$\|\boldsymbol{\phi}\|_{-1,h} = \sup_{\boldsymbol{\psi} \in \boldsymbol{V}_h} \frac{(\boldsymbol{\phi}, \boldsymbol{\psi})}{\|\boldsymbol{\psi}\|_1},$$

$$\|\omega\|_{1,h} = \|\mathbf{grad}_h \omega\|_0, \qquad \|\omega\|_{-1,h} = \sup_{\mu \in W_h} \frac{(\omega, \mu)}{\|\mu\|_{1,h}},$$

$$\|\boldsymbol{\zeta}\|_{\mathring{\boldsymbol{H}}_h(\operatorname{rot}_h)} = \left(\|\boldsymbol{\zeta}\|_0^2 + \|\operatorname{rot}_h \boldsymbol{\zeta}\|_0^2\right)^{1/2}, \qquad \|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\operatorname{div}_h)} = \sup_{\boldsymbol{\eta} \in \boldsymbol{\Gamma}_h} \frac{(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\operatorname{rot}_h)}},$$

$$\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\operatorname{div}_h) \cap t \cdot \boldsymbol{L}^2} = \left(\|\boldsymbol{\zeta}\|^2_{\boldsymbol{H}_h^{-1}(\operatorname{div}_h)} + t^2 \|\boldsymbol{\zeta}\|_0^2\right)^{1/2},$$

$$\|\boldsymbol{\zeta}\|_{\mathring{\boldsymbol{H}}_h(\operatorname{rot}_h) + t^{-1} \cdot \boldsymbol{L}^2} = \sup_{\boldsymbol{\eta} \in \boldsymbol{\Gamma}_h} \frac{(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\|\boldsymbol{\eta}\|_{\boldsymbol{H}_h^{-1}(\operatorname{div}_h) \cap t \cdot \boldsymbol{L}^2}} = \inf_{\substack{\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2 \in \boldsymbol{\Gamma}_h \\ \boldsymbol{\zeta}_1 + \boldsymbol{\zeta}_2 = \boldsymbol{\zeta}}} (\|\boldsymbol{\zeta}_1\|^2_{\mathring{\boldsymbol{H}}_h(\operatorname{rot}_h)} + t^{-2} \|\boldsymbol{\zeta}_2\|_0^2)^{1/2}.$$

We can now define the t-dependent norm on $X_{t,h} = \boldsymbol{V}_h \times W_h \times \boldsymbol{\Gamma}_h$:

$$\|(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta})\|^2_{X_{t,h}} = \|\boldsymbol{\phi}\|_1^2 + \|\omega\|_{1,h}^2 + \|\boldsymbol{\zeta}\|^2_{\boldsymbol{H}_h^{-1}(\operatorname{div}_h) \cap t \cdot \boldsymbol{L}^2}.$$

The dual space $X^*_{t,h}$ coincides with $X_{t,h}$ as a set, but carries the dual norm:

$$\|(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta})\|^2_{X^*_{t,h}} = \|\boldsymbol{\phi}\|^2_{-1,h} + \|\omega\|^2_{-1,h} + \|\boldsymbol{\zeta}\|^2_{\mathring{\boldsymbol{H}}_h(\operatorname{rot}_h) + t^{-1} \cdot \boldsymbol{L}^2}.$$

We next discuss several choices of finite element spaces which have been proposed for the approximation of the Reissner–Mindlin model and verify that they satisfy hypotheses (H1)–(H3). Although many other possibilities appear in the literature, we confine ourselves to methods using triangular finite elements which have been rigorously established to be free of locking. In the method of Arnold and Falk [1], each component of $\boldsymbol{V}_h$ consists of continuous, piecewise linear functions plus cubic bubbles, W_h is the nonconforming piecewise linear approximation of $\mathring{H}^1$, $\boldsymbol{\Gamma}_h$ is the space of piecewise constants, Q_h is the space of continuous piecewise linear functions, $\mathbf{grad}_h$ is the piecewise gradient (which is one-to-one on W_h), $\mathbf{curl}_h$ is the ordinary $\mathbf{curl}$, and $\boldsymbol{R}_h$ is the L^2-projection. The discrete Helmholtz decomposition (H1) is proven in [1]. Clearly $\|\boldsymbol{R}_h \boldsymbol{\psi}\|_0 \leq \|\boldsymbol{\psi}\|_1$. In addition

$$(\operatorname{rot}_h \boldsymbol{R}_h \boldsymbol{\psi}, q) = (\boldsymbol{R}_h \boldsymbol{\psi}, \mathbf{curl}_h q) = (\boldsymbol{\psi}, \mathbf{curl}_h q) = (\boldsymbol{\psi}, \mathbf{curl}\, q) = (\operatorname{rot} \boldsymbol{\psi}, q) \quad \text{for all } q \in Q_h.$$

Hence, $\|\operatorname{rot}_h \boldsymbol{R}_h \boldsymbol{\psi}\|_0 \leq \|\operatorname{rot} \boldsymbol{\psi}\|_0$, which establishes (H2). Note that in this case, the ordinary divergence and rotation operators are not defined on $\boldsymbol{\Gamma}_h$, which accounts for the introduction of the discrete versions of these operators. Finally, we observe that

$$(\mathbf{curl}_h p, \boldsymbol{R}_h \boldsymbol{\psi}) = (\mathbf{curl}\, p, \boldsymbol{\psi}) \quad \text{for all } p \in Q_h, \boldsymbol{\psi} \in \boldsymbol{V}_h,$$

and so (H3) follows by a simple modification of the stability proof for the MINI element for the Stokes problem.

Several families of locking-free methods are proposed and analysed in [9]. These methods fall into our framework, but have some additional properties. First, $W_h \subset \mathring{H}^1$, $\boldsymbol{\Gamma}_h \subset \mathring{\boldsymbol{H}}(\mathrm{rot})$, $\mathbf{grad}_h = \mathbf{grad}$, and $\mathrm{rot}_h = \mathrm{rot}$ (then div_h and $\mathbf{curl}_h$ are defined by duality with respect to these; they do not coincide with the ordinary divergence and curl). Second, the operator $\boldsymbol{R}_h$ extends to a bounded operator from $\boldsymbol{H}^1$ to $\boldsymbol{\Gamma}_h$ and its norm in $\mathcal{L}(\boldsymbol{H}^1, \boldsymbol{L}^2)$ is bounded uniformly with respect to h. Most important are the following five properties, which were formulated in [9] as the basis of the convergence theory.

(P1) $\mathbf{grad}\, W_h \subset \boldsymbol{\Gamma}_h$;

(P2) $\mathrm{rot}\, \boldsymbol{\Gamma}_h \subset Q_h$;

(P3) $\mathrm{rot}\, \boldsymbol{R}_h \boldsymbol{\psi} = P_h \,\mathrm{rot}\, \boldsymbol{\psi}$ for $\boldsymbol{\psi} \in \mathring{\boldsymbol{H}}^1$, where P_h is the L^2-projection into Q_h;

(P4) If $\boldsymbol{\eta} \in \boldsymbol{\Gamma}_h$ satisfies $\mathrm{rot}\, \boldsymbol{\eta} = 0$, then $\boldsymbol{\eta} = \mathbf{grad}\, \mu$ for some $\mu \in W_h$;

(P5) There is a positive constant C independent of h such that

$$\sup_{\boldsymbol{\psi} \in \boldsymbol{V}_h} \frac{(\mathrm{rot}\, \boldsymbol{\psi}, q)}{\|\boldsymbol{\psi}\|_1} \geq C \|q\|_0, \quad \text{for all } q \in Q_h.$$

These properties imply our hypotheses (H1) through (H3). The discrete Helmholtz decomposition (H1) is derived from (P1) through (P4) in [9]. Hypothesis (H2) clearly follows from the assumption that $\boldsymbol{R}_h$ is uniformly bounded in $\mathcal{L}(\boldsymbol{H}^1, \boldsymbol{L}^2)$ and (P3). Hypothesis (H3) is an easy consequence of (P3) and (P5). Consequently, our hypotheses are satisfied by all the elements proposed in [9]. In addition to those elements, Durán and Liberman [12] have proposed an element which possesses the same properties (in particular which satisfies (P1) through (P5)), and consequently which satisfies our hypotheses as well. For this element, the space W_h is chosen to be continuous, piecewise linear polynomials, $\boldsymbol{\Gamma}_h$ the lowest order Raviart-Thomas approximation of $\mathring{\boldsymbol{H}}(\mathrm{rot})$, and $\boldsymbol{R}_h$ is the usual interpolation operator associated with this space. The space $\boldsymbol{V}_h$ consists of continuous, piecewise linear polynomial vectors augmented by quadratic functions which have support in two triangles and vanish on all edges except the common edge, where their direction is along the tangent to the side. This space, together with piecewise constants, is a modification (by rotation) of a well-known stable Stokes element and hence (H3) is easily seen to be satisfied. A variant is to take for $\boldsymbol{V}_h$ all continuous, piecewise quadratic vectors.

4. The discrete isomorphism result. We now turn to the discrete isomorphism result.

THEOREM 4.1. *Suppose that the subspaces $\boldsymbol{V}_h$, W_h, $\boldsymbol{\Gamma}_h$, and Q_h and the operators $\mathbf{grad}_h$, $\mathbf{curl}_h$, and $\boldsymbol{R}_h$ satisfy hypotheses (H1)–(H3) and let $(\boldsymbol{f}_h, g_h, \boldsymbol{j}_h) \in X_{t,h}$ be given. Then there is a unique solution $(\boldsymbol{\phi}_h, \omega_h, \boldsymbol{\zeta}_h) \in X_{t,h}$ to Problem (G_h). Moreover there exist positive constants c_1 and c_2, independent of h and t, such that for $0 \leq t \leq 1$,*

$$c_1 \|(\boldsymbol{f}_h, g_h, \boldsymbol{j}_h)\|_{X^*_{t,h}} \leq \|(\boldsymbol{\phi}_h, \omega_h, \boldsymbol{\zeta}_h)\|_{X_{t,h}} \leq c_2 \|(\boldsymbol{f}_h, g_h, \boldsymbol{j}_h)\|_{X^*_{t,h}}.$$

Before turning to the proof, we prove two lemmas under the hypotheses (H1)–(H3). The first lemma uses the discrete Helmholtz decomposition to give an equivalent norm on $\boldsymbol{H}_h^{-1}(\mathrm{div}_h)$.

Lemma 4.2. *Let $\boldsymbol{\zeta} = \mathbf{grad}_h\, r + \mathbf{curl}_h\, p$, with $r \in W_h$ and $p \in Q_h$. Then there exists a constant $c > 0$, independent of h such that*

$$\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)} \le \left(\|\,\mathbf{grad}_h\, r\|_0^2 + \|p\|_0^2\right)^{1/2} \le c\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)}.$$

Proof. The first inequality is straightforward:

$$\begin{aligned}\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)} &= \sup_{\boldsymbol{\eta}\in\boldsymbol{\Gamma}_h} \frac{(\boldsymbol{\zeta},\boldsymbol{\eta})}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}} = \sup_{\boldsymbol{\eta}\in\boldsymbol{\Gamma}_h} \frac{[(\mathbf{grad}_h\, r,\boldsymbol{\eta}) + (p, \mathrm{rot}_h\,\boldsymbol{\eta})]}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}} \\ &\le \sup_{\boldsymbol{\eta}\in\boldsymbol{\Gamma}_h} \frac{[\|\,\mathbf{grad}_h\, r\|_0\|\boldsymbol{\eta}\|_0 + \|p\|_0\|\,\mathrm{rot}_h\,\boldsymbol{\eta}\|_0]}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}} \le \left(\|\,\mathbf{grad}_h\, r\|_0^2 + \|p\|_0^2\right)^{1/2}.\end{aligned}$$

To prove the second inequality, choose

$$\boldsymbol{\eta} = \mathbf{grad}_h\, r + \mathbf{curl}_h\, q,$$

where $q \in Q_h$ satisfies

$$(\mathbf{curl}_h\, q, \mathbf{curl}_h\, v) = (p, v) \quad \text{for all } v \in Q_h. \tag{4.1}$$

Note that this problem, which in light of (3.6) has a unique solution, is equivalent to

$$\mathrm{rot}_h\,\mathbf{curl}\, q_h = p.$$

From (4.1) and (3.6) we have $\|\,\mathbf{curl}_h\, q\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} \le (1 + C^2)^{1/2}\|p\|_0$, so

$$\|\boldsymbol{\eta}\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} = \|\,\mathbf{grad}_h\, r\|_0^2 + \|\,\mathbf{curl}_h\, q\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} \le \|\,\mathbf{grad}_h\, r\|_0^2 + C\|p\|_0^2.$$

Using the orthogonality of $\mathbf{grad}_h\, W_h$ and $\mathbf{curl}_h\, Q_h$, we get

$$\begin{aligned}\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)} \ge \frac{(\boldsymbol{\zeta},\boldsymbol{\eta})}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}} &\ge \frac{(\mathbf{grad}_h\, r, \mathbf{grad}_h\, r) + (\mathbf{curl}_h\, p, \mathbf{curl}_h\, q)}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}} \\ &= \frac{\|\,\mathbf{grad}_h\, r\|_0^2 + \|p\|_0^2}{\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}},\end{aligned}$$

and the desired result easily follows. □

Lemma 4.3. *There exists a constant $\beta > 0$ independent of h such that for all $\boldsymbol{\zeta} \in \boldsymbol{\Gamma}_h$,*

$$\sup_{\boldsymbol{\psi}\in V_h, \mu\in W_h} \frac{(\boldsymbol{\zeta}, \boldsymbol{R}_h\boldsymbol{\psi} - \mathbf{grad}_h\, \mu)}{(\|\boldsymbol{\psi}\|_1^2 + \|\mathbf{grad}_h\, \mu\|_0^2)^{1/2}} \geq \beta \|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)}.$$

Proof. By (H1), we can write $\boldsymbol{\zeta} = \mathbf{grad}_h\, r + \mathbf{curl}_h\, p$ for some $r \in W_h$, $p \in Q_h$. By (H3), we can then find $\boldsymbol{\psi} \in \boldsymbol{V}_h$ such that

$$(\mathbf{curl}_h\, p, \boldsymbol{R}_h\boldsymbol{\psi}) = \|p\|_0^2, \quad \|\boldsymbol{\psi}\|_1 \leq C_2\|p\|_0.$$

Next, choose $\mu = -C_1^2C_2^2 r$, where C_1 and C_2 are the constants in (H2) and (H3). Then

$$\begin{aligned}(\boldsymbol{\zeta}, \boldsymbol{R}_h\boldsymbol{\psi} - \mathbf{grad}_h\, \mu) &= (\mathbf{curl}_h\, p, \boldsymbol{R}_h\boldsymbol{\psi}) - (\mathbf{grad}_h\, r, \mathbf{grad}_h\, \mu) + (\mathbf{grad}_h\, r, \boldsymbol{R}_h\boldsymbol{\psi})\\ &\geq \|p\|_0^2 + C_1^2C_2^2\|\mathbf{grad}_h\, r\|_0^2 - \|\mathbf{grad}_h\, r\|_0\|\boldsymbol{R}_h\boldsymbol{\psi}\|_0\\ &\geq \|p\|_0^2 + C_1^2C_2^2\|\mathbf{grad}_h\, r\|_0^2 - C_1C_2\|\mathbf{grad}_h\, r\|_0\|p\|_0\\ &\geq \frac{1}{2}\|p\|_0^2 + \frac{C_1^2C_2^2}{2}\|\mathbf{grad}_h\, r\|_0^2.\end{aligned}$$

Moreover

$$\|\boldsymbol{\psi}\|_1^2 + \|\mathbf{grad}_h\, \mu\|_0^2 \leq C(\|p\|_0^2 + \|\mathbf{grad}_h\, r\|_0^2).$$

The result follows from these two inequalities and the proceeding lemma. ∎

Proof of Theorem 4.1. For simplicity of notation we drop the subscript h on the solution and data, writing them as $(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta})$ and $(\boldsymbol{f}, g, \boldsymbol{j})$ respectively. We first prove the bound on the solution from above:

$$\|(\boldsymbol{\phi}, \omega, \boldsymbol{\zeta}\|_{X_{t,h}} \leq c_2\|(\boldsymbol{f}, g, \boldsymbol{j})\|_{X_{t,h}^*}. \tag{4.2}$$

Choosing $\boldsymbol{\psi} = \boldsymbol{\phi}$, $\mu = \omega$, and $\boldsymbol{\eta} = \boldsymbol{\zeta}$ in (3.1)–(3.3), we get

$$\begin{aligned}&(\mathcal{C}\,\mathcal{E}\boldsymbol{\phi}, \mathcal{E}\,\boldsymbol{\phi}) + t^2(\boldsymbol{\zeta}, \boldsymbol{\zeta}) = (\boldsymbol{f}, \boldsymbol{\phi}) + (g, \omega) - (\boldsymbol{j}, \boldsymbol{\zeta})\\ &\qquad \leq \|\boldsymbol{f}\|_{-1,h}\|\boldsymbol{\phi}\|_1 + \|g\|_{-1,h}\|\omega\|_{1,h} + \|\boldsymbol{j}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2}\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2}.\end{aligned} \tag{4.3}$$

From Lemma 4.3, (3.1), and (3.2), we have

$$\begin{aligned}\beta\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)} &\leq \sup_{\boldsymbol{\psi}\in V_h, \mu\in W_h} \frac{(\boldsymbol{\zeta}, \boldsymbol{R}_h\boldsymbol{\psi} - \mathbf{grad}_h\, \mu)}{(\|\boldsymbol{\psi}\|_1^2 + \|\mathbf{grad}_h\, \mu\|_0^2)^{1/2}}\\ &\leq \sup_{\boldsymbol{\psi}\in V_h, \mu\in W_h} \frac{[(\mathcal{C}\,\mathcal{E}\,\boldsymbol{\phi}, \mathcal{E}\,\boldsymbol{\psi}) - (\boldsymbol{f}, \boldsymbol{\psi}) - (g, \mu)]}{(\|\boldsymbol{\psi}\|_1^2 + \|\mathbf{grad}_h\, \mu\|_0^2)^{1/2}}\\ &\leq C[\|\boldsymbol{\phi}\|_1 + \|\boldsymbol{f}\|_{-1,h} + \|g\|_{-1,h}].\end{aligned} \tag{4.4}$$

From (3.3), we have

$$\mathbf{grad}_h\,\omega = \boldsymbol{R}_h\boldsymbol{\phi} + t^2\boldsymbol{\zeta} + \boldsymbol{j}. \tag{4.5}$$

Now

$$\|\boldsymbol{R}_h\boldsymbol{\phi}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} \le \|\boldsymbol{R}_h\boldsymbol{\phi}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} \le C\|\boldsymbol{\phi}\|_1 \tag{4.6}$$

using the definition of the sum norm and hypothesis (H2). It also follows from the definition of the norms that

$$t^2\|\boldsymbol{\zeta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} \le t\|\boldsymbol{\zeta}\|_0 \le \|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2}. \tag{4.7}$$

Combining (4.5), (4.6), and (4.7) we deduce

$$\|\,\mathbf{grad}_h\,\omega\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} \le C[\|\boldsymbol{\phi}\|_1 + \|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2} + \|\boldsymbol{j}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2)}].$$

Now by the definition of the norm and (H1),

$$\begin{aligned}
\|\,\mathbf{grad}_h\,\omega\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} &= \inf_{\substack{\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2\in\boldsymbol{\Gamma}_h\\ \boldsymbol{\zeta}_1+\boldsymbol{\zeta}_2=\mathbf{grad}_h\,\omega}} (\|\boldsymbol{\zeta}_1\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} + t^{-2}\|\boldsymbol{\zeta}_2\|_0^2)\\
&= \inf_{r\in W_h, p\in Q_h} (\|\,\mathbf{grad}_h(\omega - r) - \mathbf{curl}_h\,p\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} + t^{-2}\|\,\mathbf{grad}_h\,r + \mathbf{curl}_h\,p\|_0^2)\\
&= \inf_{r\in W_h, p\in Q_h} (\|\,\mathbf{grad}_h(\omega - r)\|_0^2 + \|\,\mathbf{curl}_h\,p\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}\\
&\qquad + t^{-2}\|\,\mathbf{grad}_h\,r\|_0^2 + t^{-2}\|\,\mathbf{curl}_h\,p\|_0^2)\\
&= \inf_{r\in W_h} (\|\,\mathbf{grad}_h(\omega - r)\|_0^2 + t^{-2}\|\,\mathbf{grad}_h\,r\|_0^2)\\
&= \frac{t^{-2}}{1+t^{-2}}\|\,\mathbf{grad}_h\,\omega\|_0^2 = \frac{1}{1+t^2}\|\,\mathbf{grad}_h\,\omega\|_0^2,
\end{aligned}$$

where the last line above is obtained by a simple variational argument which shows that the infimum is obtained for $r = \omega/(1+t^{-2})$.

Since we have assumed that $t \le 1$, it easily follows that

$$\frac{1}{\sqrt{2}}\|\omega\|_{1,h} \le \|\,\mathbf{grad}_h\,\omega\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} \le \|\omega\|_{1,h}, \tag{4.8}$$

and so

$$\|\omega\|_{1,h} \le C(\|\boldsymbol{\phi}\|_1 + \|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2} + \|\boldsymbol{j}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2}). \tag{4.9}$$

Inserting this result in (4.3) and using the Schwarz and arithmetic-geometric mean inequalities, we obtain

$$\|\phi\|_1^2 + t^2\|\zeta\|_0^2 \le C[\|f\|_{-1,h}^2 + \|g\|_{-1,h}^2 + \|j\|^2_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2} + (\|g\|_{-1,h} + \|j\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2})\|\zeta\|_{H_h^{-1}(\mathrm{div}_h)\cap t\cdot L^2}]$$

and hence that

$$\|\phi\|_1^2 + t^2\|\zeta\|_0^2 \le C[\|f\|_{-1,h}^2 + \|g\|_{-1,h}^2 + \|j\|^2_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2} + (\|g\|_{-1,h} + \|j\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2})\|\zeta\|_{H_h^{-1}(\mathrm{div}_h)}].$$

From (4.4), it then follows by standard estimates that

$$\|\phi\|_1^2 + \|\zeta\|^2_{H_h^{-1}(\mathrm{div}_h)\cap t\cdot L^2} \le C[\|f\|_{-1,h}^2 + \|g\|_{-1,h}^2 + \|j\|^2_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2}].$$

This estimate together with (4.9) completes the proof of (4.2).

The proof of the reverse bound

$$c_1\|(f, g, j)\|_{X_{t,h}^*} \le \|(\phi, \omega, \zeta\|_{X_{t,h}},$$

is quite direct. We see from (3.1) and (4.6) that

$$\begin{aligned}\|f\|_{-1,h} &= \sup_{\psi\in V_h} \frac{(\mathcal{C}\,\mathcal{E}\,\phi, \mathcal{E}\,\psi) - (\zeta, R_h\psi)}{\|\psi\|_1} \\ &\le \sup_{\psi\in V_h} \frac{[C\|\phi\|_1\|\psi\|_1 + \|\zeta\|_{H_h^{-1}(\mathrm{div}_h)\cap t\cdot L^2}\|R_h\psi\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2}]}{\|\psi\|_1} \\ &\le C(\|\phi\|_1 + \|\zeta\|_{H_h^{-1}(\mathrm{div}_h)\cap t\cdot L^2}).\end{aligned}$$

Similarly, (3.2) and (4.8) give

$$\begin{aligned}\|g\|_{-1,h} &= \sup_{\mu\in W_h} \frac{(\zeta, \mathbf{grad}_h\,\mu)}{\|\mu\|_{1,h}} \\ &\le \sup_{\mu\in W_h} \frac{(\zeta, \mathbf{grad}_h\,\mu)}{\|\,\mathbf{grad}_h\,\mu\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2}} \\ &\le \|\zeta\|_{H_h^{-1}(\mathrm{div}_h)\cap t\cdot L^2}.\end{aligned}$$

Finally, from (4.5), (4.6), (4.8), and (4.7), we obtain

$$\begin{aligned}\|j\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2} &\le \|R_h\phi\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2} + \|\,\mathbf{grad}_h\,\omega\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2} \\ &\quad + t^2\|\zeta\|_{\mathring{H}_h(\mathrm{rot}_h)+t^{-1}\cdot L^2} \\ &\le C(\|\phi\|_1 + \|\omega\|_{1,h} + \|\zeta\|_{H_h^{-1}(\mathrm{div}_h)\cap t\cdot L^2}).\end{aligned}$$

The desired inequality follows by combining these results. ∎

5. Implications for preconditioning. As discussed in the introduction, we propose to solve the discrete system (3.4) by applying the preconditioned minimum residual method (or another iterative scheme with similar properties). Furthermore, we will use Theorem 4.1 to derive the desired properties of the preconditioner. The preconditioner $\mathcal{B}_{t,h} : X_h \to X_h$ is required to be L^2-symmetric and positive definite. Hence, since $\mathcal{A}_{t,h}$ is also L^2-symmetric, the preconditioned linear system

$$\mathcal{B}_{t,h}\mathcal{A}_{t,h}\begin{pmatrix}\boldsymbol{\phi}_h\\ \omega_h\\ \boldsymbol{\zeta}_h\end{pmatrix} = \mathcal{B}_{t,h}\begin{pmatrix}\boldsymbol{f}_h\\ g_h\\ \boldsymbol{j}_h\end{pmatrix} \tag{5.1}$$

has a coefficient operator which is symmetric with respect to the inner product $(\mathcal{B}_{t,h}^{-1}\,\cdot\,,\,\cdot\,)$. The preconditioned minimum residual method is a Krylov space method where at least one evaluation of the coefficient operator $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$ is necessary for each iteration. However, in order to improve the numerical stability of the method, an iteration which requires two evaluations of the original coefficient operator $\mathcal{A}_{t,h}$ and one evaluation of the preconditioner $\mathcal{B}_{t,h}$ is often preferred. For discussions on implementations of the preconditioned minimum residual method we refer to Wathen and Silvester [24], Klawonn [16], and Chapter 9 of the text [14] by Hackbusch.

The preconditioned minimum residual method gives an optimal approximation of the solution of the linear system in the Krylov space generated by the operator $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$ in the norm associated with the inner product

$$(\mathcal{B}_{t,h}\mathcal{A}_{t,h}\,\cdot\,,\mathcal{A}_{t,h}\,\cdot\,). \tag{5.2}$$

From this optimality property one can easily derive upper bounds for the error with respect to spectral properties of the operator $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$ (cf. for example [19], [24], or [14]). Let $\kappa = \kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ be the spectral condition number given by

$$\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h}) = \frac{\sup|\lambda|}{\inf|\lambda|},$$

where the supremum and infimum is taken over the spectrum of $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$. Then the reduction factor, after n iterations, in the norm associated with the inner product (5.2), is bounded by $2r^n/(1+r^{2n})$, where $r^2 = (\kappa-1)/(\kappa+1)$. This upper bound for the reduction factor is in fact the same as one would obtain after $n/2$ iterations of the conjugate gradient method applied to the normal equations associated with the preconditioned system (5.1). However, as for example discussed in [5] and [19], the minimum residual method will usually perform much better than a normal equation approach, and it is well understood that this phenomenon can be explained from the lack of symmetry around the origin of the spectrum of $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$.

However, the difference in the performance of these two methods is not important for the theoretical discussion given here. The significance of the upper bound given above, is that if the spectral condition number of $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$ is bounded independently of t and h, the resulting iteration will achieve any given reduction factor in a finite number of iterations, independent of t and h.

As indicated above, we will use Theorem 4.1 in order to derive the desired properties of the preconditioner $\mathcal{B}_{t,h}$. This theorem states that the operator norms $\|\mathcal{A}_{t,h}\|_{\mathcal{L}(X_{t,h},X^*_{t,h})}$ and $\|\mathcal{A}^{-1}_{t,h}\|_{\mathcal{L}(X^*_{t,h},X_{t,h})}$ are bounded independently of t and h. Hence, if the preconditioner $\mathcal{B}_{t,h}$ is chosen such that $\|\mathcal{B}_{t,h}\|_{\mathcal{L}(X^*_{t,h},X_{t,h})}$ and $\|\mathcal{B}^{-1}_{t,h}\|_{\mathcal{L}(X_{t,h},X^*_{t,h})}$ also are bounded independently of t and h, we immediately obtain bounds, independent of t and h, for the operator $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$, and its inverse in $\mathcal{L}(X_{t,h}, X_{t,h})$. Since the spectral radius of an operator is bounded by any operator norm (of an operator mapping a space into itself), we therefore conclude that this mapping property of $\mathcal{B}_{t,h}$ implies that the spectral condition number of $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$ is bounded independently of t and h.

Assume that we have used the mapping argument above to identify one possible preconditioner $\mathcal{B}_{t,h}$, such that the operator norms $\|\mathcal{B}_{t,h}\|_{\mathcal{L}(X^*_{t,h},X_{t,h})}$ and $\|\mathcal{B}^{-1}_{t,h}\|_{\mathcal{L}(X_{t,h},X^*_{t,h})}$ are bounded uniformly in t and h. Furthermore, assume that $\bar{\mathcal{B}}_{t,h} : X_h \to X_h$ is another L^2-symmetric operator which is spectrally equivalent to $\mathcal{B}_{t,h}$, i.e., the two bilinear forms $(\mathcal{B}_{t,h}\cdot,\cdot)$ and $(\bar{\mathcal{B}}_{t,h}\cdot,\cdot)$ are equivalent uniformly in t and h. Since $\|\cdot\|_{X_{t,h}}$ and $\|\cdot\|_{X^*_{t,h}}$ are dual norms with respect to the L^2 inner product we can conclude that $\|\bar{\mathcal{B}}_{t,h}\|_{\mathcal{L}(X^*_{t,h},X_{t,h})}$ and $\|\bar{\mathcal{B}}^{-1}_{t,h}\|_{\mathcal{L}(X_{t,h},X^*_{t,h})}$ are also bounded independently of t and h. Hence, we can always replace one possible preconditioner by another L^2-symmetric and spectrally equivalent operator.

Utilizing the structure of the space X_h as a product space, we consider block diagonal, positive definite preconditioners of the form

$$\mathcal{B}_{t,h} = \begin{pmatrix} \boldsymbol{L}_h & 0 & 0 \\ 0 & M_h & 0 \\ 0 & 0 & \boldsymbol{N}_{t,h} \end{pmatrix} : X_{t,h} \to X_{t,h}. \tag{5.3}$$

Written in terms of the blocks of $\mathcal{B}_{t,h}$, the desired mapping properties of the preconditioner are equivalent to the existence of constants c_1 and c_2, independent of t and h such that

$$c_1\|\boldsymbol{f}_h\|_{-1,h} \le \|\boldsymbol{L}_h\boldsymbol{f}_h\|_1 \le c_2\|\boldsymbol{f}_h\|_{-1,h}, \tag{5.4}$$

$$c_1\|g_h\|_{-1,h} \le \|M_h g_h\|_{1,h} \le c_2\|g_h\|_{-1,h}, \tag{5.5}$$

$$c_1\|\boldsymbol{j}_h\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} \le \|\boldsymbol{N}_{t,h}\boldsymbol{j}_h\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2} \le c_2\|\boldsymbol{j}_h\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2}. \tag{5.6}$$

The conditions on $\boldsymbol{L}_h$ and M_h required for (5.4) and (5.5) are exactly of the type satisfied by standard preconditioners for second order elliptic operators. Hence, we need only discuss the operator $\boldsymbol{N}_{t,h}$. The special case $t = 0$ was discussed in § 3 for the continuous problem, where it was noted that in this case $\boldsymbol{N}_0$ is an isomorphism from $\mathring{\boldsymbol{H}}(\mathrm{rot})$ into $\boldsymbol{H}^{-1}(\mathrm{div})$ and that the natural choice for $\boldsymbol{N}_0$ is given by $\boldsymbol{I} + \mathbf{curl}\,\mathrm{rot}$. For the discrete problem, an analogous situation holds: if $\boldsymbol{N}_{0,h}$ is chosen to be the operator $\boldsymbol{\Lambda}_h : \boldsymbol{\Gamma}_h \to \boldsymbol{\Gamma}_h$ defined by $\boldsymbol{\Lambda}_h = \boldsymbol{I} + \mathbf{curl}_h\,\mathrm{rot}_h$, then (5.6) holds in the case $t = 0$ with $c_1 = c_2 = 1$. In fact, if we choose $\boldsymbol{N}_{0,h}$ to be an operator which is spectrally equivalent to $\boldsymbol{I} + \mathbf{curl}_h\,\mathrm{rot}_h$ then $\boldsymbol{N}_{0,h}$ also has the right mapping property in this case.

We now consider the case of a general t. Then (5.6) tells us that $\boldsymbol{N}_{t,h}$ should be spectrally equivalent to an isomorphism from $\boldsymbol{\Gamma}_h$ equipped with the norm $\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2$ to the same space equipped with the norm $\boldsymbol{H}_h^{-1}(\mathrm{div}_h) \cap t \cdot \boldsymbol{L}^2$. Therefore the following result is useful.

Lemma 5.1. *The operator $\boldsymbol{\Lambda}_h^{-1} + t^2\boldsymbol{I}$ maps $\boldsymbol{\Gamma}_h$ isomorphically onto itself. Moreover*

$$\|(\boldsymbol{\Lambda}_h^{-1} + t^2\boldsymbol{I})\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} = \|\boldsymbol{\eta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2} \quad \textit{for all } \boldsymbol{\eta} \in \boldsymbol{\Gamma}_h.$$

Proof. Indeed,

$$((\boldsymbol{\Lambda}_h^{-1} + t^2\boldsymbol{I})\boldsymbol{\eta}, \boldsymbol{\eta}) = \|\boldsymbol{\eta}\|^2_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2}.$$

The lemma follows from this identity and the fact that the $\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h) + t^{-1} \cdot \boldsymbol{L}^2$ norm is dual to the $\boldsymbol{H}_h^{-1}(\mathrm{div}_h) \cap t \cdot \boldsymbol{L}^2$ norm. □

Having found an operator which gives the correct mapping properties, the problem is to find an equivalent operator which is easy to apply. Clearly, we do not want to apply $(\boldsymbol{\Lambda}_h^{-1} + t^2\boldsymbol{I})^{-1}$. For this purpose, it is convenient to introduce a family of operators on $\boldsymbol{\Gamma}_h$ depending on the parameter t. For each $t \in [0,1]$ define the operator $\boldsymbol{\Lambda}_{t,h} : \boldsymbol{\Gamma}_h \to \boldsymbol{\Gamma}_h$ by

$$\boldsymbol{\Lambda}_{t,h} = \boldsymbol{I} + t^2\,\mathbf{curl}_h\,\mathrm{rot}_h,$$

so that $\boldsymbol{\Lambda}_h = \boldsymbol{\Lambda}_{1,h}$. The lemma above shows that the operator $(\boldsymbol{\Lambda}_h^{-1} + t^2\boldsymbol{I})^{-1} = \boldsymbol{\Lambda}_h(\boldsymbol{I} + t^2\boldsymbol{\Lambda}_h)^{-1}$ has the mapping property required by (5.6).

Observe that since $t \in [0,1]$, the operator $(\boldsymbol{\Lambda}_h^{-1} + t^2\boldsymbol{I})^{-1}$ is spectrally equivalent to $\boldsymbol{\Lambda}_{1,h}\boldsymbol{\Lambda}_{t,h}^{-1}$. To see this, note that if $\boldsymbol{\eta}$ is an eigenfunction of $\boldsymbol{\Lambda}_h$ with eigenvalue μ, then $\boldsymbol{\eta}$ is also an eigenfunction of $(\boldsymbol{\Lambda}_h^{-1} + t^2\boldsymbol{I})^{-1}$ and $\boldsymbol{\Lambda}_{1,h}\boldsymbol{\Lambda}_{t,h}^{-1}$ with eigenvalues $\mu/(1+\mu t^2)$ and $\mu/(1+\mu t^2 - t^2)$, respectively. Since $\mu > 0$, we see immediately that for $0 \le t \le 1$, $\mu/(1+\mu t^2) \le \mu/(1+\mu t^2 - t^2)$. To show that

$$\mu/(1+\mu t^2 - t^2) \le c\mu/(1+\mu t^2),$$

we show that

$$f(t,\mu) = (1+\mu t^2)/(1+\mu t^2 - t^2) \le c.$$

Now since $f(t,\mu)$ is an increasing function of t and a decreasing function of μ,

$$f(t,\mu) \le f(1,\mu_0) = (1+\mu_0)/\mu_0,$$

where μ_0 is the minimum eigenvalue of $\mathbf{\Lambda}_h$. Since $(\mathbf{\Lambda}_h \boldsymbol{\zeta}, \boldsymbol{\zeta}) \ge (\boldsymbol{\zeta}, \boldsymbol{\zeta})$, $\mu_0 \ge 1$ and so $f(t,\mu) \le 2$.

Hence, it seems natural to derive the preconditoner $\boldsymbol{N}_{t,h}$ by replacing $\mathbf{\Lambda}_{t,h}^{-1}$ by a suitable preconditioner. However, such an approach will lead to an operator $\boldsymbol{N}_{t,h}$ which, in general, will not be L^2-symmetric. Instead we shall derive the preconditioner from the identity

$$\mathbf{\Lambda}_{1,h}\mathbf{\Lambda}_{t,h}^{-1} = \boldsymbol{I} + (1-t^2)\,\mathbf{curl}_h\,\mathrm{rot}_h\,\mathbf{\Lambda}_{t,h}^{-1}. \tag{5.7}$$

In order to use this identity, we define operators $S_{t,h} : Q_h \to Q_h$ by

$$S_{t,h} = I + t^2\,\mathrm{rot}_h\,\mathbf{curl}_h\,.$$

These operators are L^2-symmetric and positive definite and correspond to finite element discretization of the Neumann problem for the scalar second order elliptic operator $I - t^2\Delta$. In fact, if $\mathrm{rot}_h = \mathrm{rot}$, the operator $S_{t,h}$ corresponds to a standard mixed finite element discretization of such problems, except that the curl and rotation operators are used instead of the gradient and divergence. On the other hand, for the method proposed in [1], $S_{t,h}$ corresponds to a standard conforming piecewise linear discretization of this operator.

Observe that from the definitions of $\mathbf{\Lambda}_{t,h}$ and $S_{t,h}$ we obviously have the identity

$$\mathrm{rot}_h\,\mathbf{\Lambda}_{t,h} = S_{t,h}\,\mathrm{rot}_h\,.$$

This immediately implies that

$$\mathrm{rot}_h\,\mathbf{\Lambda}_{t,h}^{-1} = S_{t,h}^{-1}\,\mathrm{rot}_h\,.$$

Hence, it follows from (5.7) that

$$\mathbf{\Lambda}_{1,h}\mathbf{\Lambda}_{t,h}^{-1} = \boldsymbol{I} + (1-t^2)\,\mathbf{curl}_h\,S_{t,h}^{-1}\,\mathrm{rot}_h\,. \tag{5.8}$$

We now assume that we have at our disposal preconditioners $\Phi_{t,h} : Q_h \to Q_h$ for the discrete elliptic operators $S_{t,h}$. More precisely, we assume that the operators $\Phi_{t,h}$ are self-adjoint operators which are spectrally equivalent to $S_{t,h}^{-1}$, i.e., there exist positive constants c_1 and c_2, independent of t and h, such that

$$c_1(\Phi_{t,h}q,q) \le (S_{t,h}^{-1}q,q) \le c_2(\Phi_{t,h}q,q) \quad \text{for all } q \in Q_h. \tag{5.9}$$

We will not consider the construction of the operators $\Phi_{t,h}$ here. But as observed above, the operators $S_{t,h}$ correspond to conforming or nonconforming finite element approximations of the elliptic operators $I - t^2\Delta$. In the conforming case, the construction of such preconditioners has been intensively studied. Also preconditioners for nonconforming approximations have been studied by many authors. We mention, for example, that techniques where such preconditioners are derived from standard conforming preconditioners are described by Bramble, Pasciak and Xu [7] and Xu [25], while studies of the particular operators which arise from mixed finite element discretizations are performed by Cowsar [10], Cowsar, Mandel and Wheeler [11], Rusten and Winther [20], Rusten, Vassilevski and Winther [18], and Vassilevski and Wang [22].

Using the operators $\Phi_{t,h}$, we define operators $\boldsymbol{D}_{t,h} : \boldsymbol{\Gamma}_h \to \boldsymbol{\Gamma}_h$ by

$$\boldsymbol{D}_{t,h} = \boldsymbol{I} + (1 - t^2)\, \mathbf{curl}_h\, \Phi_{t,h}\, \mathrm{rot}_h\, . \tag{5.10}$$

By construction these operators are L^2-symmetric. Using this operator, we can now state the main result of this section.

Theorem 5.2. *Assume that the operators $\Phi_{t,h}$ satisfy* (5.9). *Then the choice $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{t,h}$ satisfies* (5.6) *with constants c_1 and c_2 independent of t and h and hence, together with standard preconditioners $\boldsymbol{L}_h$ and M_h satisfying* (5.4) *and* (5.5) *give rise to a preconditioner $\mathcal{B}_{t,h}$ for the Reissner–Mindlin system* (3.4) *such that the spectral radius of $\mathcal{A}_{t,h}\mathcal{B}_{t,h}$ is bounded independent of t and h.*

Proof. In light of the previous discussion, it is enough to show that $\boldsymbol{D}_{t,h}$ is spectrally equivalent to the operator $\boldsymbol{\Lambda}_{1,h}\boldsymbol{\Lambda}_{t,h}^{-1}$ uniformly in t and h. This follows directly from (5.8) and (5.10). ∎

There are two cases in which the computation of the operator $\boldsymbol{D}_{t,h}$ is simplified. These are $t = 0$ for which $S_{t,h} = I$ and hence $\Phi_{t,h}$ is an operator equivalent to the identity and $t = 1$ for which $\boldsymbol{D}_{t,h} = \boldsymbol{I}$. We now discuss conditions under which the preconditioner $\boldsymbol{D}_{t,h}$ may be replaced by the simpler preconditioner $\boldsymbol{D}_{0,h}$ or $\boldsymbol{D}_{1,h} = \boldsymbol{I}$ without significant change in the convergence properties of the iteration scheme. We first consider the case when $t = O(h)$. The following lemma is the key to our result.

Lemma 5.3. *Suppose that there exists a constant K_1 such that $\|\mathrm{rot}_h\, \boldsymbol{\eta}\|_0 \le K_1 h^{-1} \|\boldsymbol{\eta}\|_0$ for all $\boldsymbol{\eta} \in \boldsymbol{\Gamma}_h$, and that we impose the restriction that $t \le K_2 h$ for another constant K_2. Then there exists a constant C depending on K_1 and K_2 but otherwise independent of h and t such that*

$$\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot \boldsymbol{L}^2} \le \|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} \le C \|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot \boldsymbol{L}^2}, \tag{5.11}$$

$$\|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)} \le \|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot \boldsymbol{L}^2} \le C \|\boldsymbol{\zeta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)}. \tag{5.12}$$

Proof. The first inequality of (5.11) is obvious. For the second note that

$$\|\boldsymbol{\eta}\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} = \|\boldsymbol{\eta}\|_0^2 + \|\,\mathrm{rot}_h\,\boldsymbol{\eta}\|_0^2 \le (1 + K_1^2 h^{-2})\|\boldsymbol{\eta}\|_0^2 \le (1 + K_1^2 K_2^2 t^{-2})\|\boldsymbol{\eta}\|_0^2 \le Ct^{-2}\|\boldsymbol{\eta}\|_0^2.$$

Therefore, for any splitting of $\boldsymbol{\zeta}$ as $\boldsymbol{\zeta}_1 + \boldsymbol{\zeta}_2$ with $\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2 \in \boldsymbol{\Gamma}_h$,

$$\|\boldsymbol{\zeta}\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} \le 2(\|\boldsymbol{\zeta}_1\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} + \|\boldsymbol{\zeta}_2\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}) \le C(\|\boldsymbol{\zeta}_1\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} + t^{-2}\|\boldsymbol{\zeta}_2\|_0^2),$$

and, taking the infimum over all such splittings,

$$\|\boldsymbol{\zeta}\|^2_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)} \le C\|\boldsymbol{\zeta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2},$$

which completes the proof of (5.11). Then (5.12) follows by duality. □

We note that when $\mathrm{rot}_h = \mathrm{rot}$, as is the case in the methods proposed in [9] and [12], the inverse hypothesis of the lemma is a standard one. For the method of [1], in which $\mathrm{rot}_h \ne \mathrm{rot}$, the verification of this hypothesis follows easily from (3.5) and the standard inverse hypothesis for piecewise linear functions.

Using this result, we see that if $t = O(h)$, then Theorem 4.1 and (5.6) remain valid when we replace the norms $\|\cdot\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2}$ and $\|\cdot\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2}$ by their values when $t = 0$, that is by $\|\cdot\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)}$ and $\|\cdot\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)}$. Thus, we have the following corollary to Theorem 5.2.

Corollary 5.4. *If $t = O(h)$, then Theorem 5.2 holds with $\boldsymbol{D}_{t,h}$ replaced by $\boldsymbol{D}_{0,h}$.*

We next consider the case when $t = O(1)$.

Corollary 5.5. *If $0 < t_0 \le t \le 1$, then Theorem 5.2 holds with $\boldsymbol{D}_{t,h}$ replaced by $\boldsymbol{D}_{1,h} = \boldsymbol{I}$, where the constant c_2 now depends on t_0, but is otherwise independent of t and h.*

Proof. We first observe that it follows easily from the definitions that for all $\boldsymbol{\eta} \in \boldsymbol{\Gamma}_h$,

$$\begin{gathered}\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+1\cdot\boldsymbol{L}^2} \le \|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2} \le \max(1, t^{-1})\|\boldsymbol{\eta}\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+1\cdot\boldsymbol{L}^2},\\ \min(1,t)\|\boldsymbol{\eta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap 1\cdot\boldsymbol{L}^2} \le \|\boldsymbol{\eta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2} \le \|\boldsymbol{\eta}\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap 1\cdot\boldsymbol{L}^2}.\end{gathered}$$

Thus, Theorem 4.1 and (5.6) remain valid when we replace the norms $\|\cdot\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2}$ and $\|\cdot\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2}$ by their values when $t = 1$, that is by $\|\cdot\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+1\cdot\boldsymbol{L}^2}$ and $\|\cdot\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap 1\cdot\boldsymbol{L}^2}$, and simultaneously replace the constant c_2 by $t^{-2}c_2$. The Corollary follows directly since $t_0 \le t$. □

The importance of this result is that the simple preconditioner resulting from the choice $\boldsymbol{N}_{t,h} = \boldsymbol{I}$ has a spectral radius independent of h and thus would be expected to produce reasonable results for moderately sized values of t. We explore this possibility in § 8 in our numerical examples.

One item that may not be clear is why in the case $t = 0$ (or in fact $t = O(h)$), we cannot just make the choice $\boldsymbol{N}_{t,h} = \boldsymbol{\Lambda}_h$. Since this is a discretization of a differential operator, intuitively it should be local, and hence easy to evaluate. The issue here, which we explore in detail in the next section, is that when one looks at the matrix representation of this operator, its application appears to require the inversion of a Gram matrix.

6. Representation of operators. The purpose of this section is to discuss the computational consequences of using the preconditioners $\boldsymbol{D}_{t,h}$ defined by (5.10). Of course, in order for these preconditioners to be effective, the cost of evaluating $\boldsymbol{D}_{t,h}$ should be proportional to the dimension of $\boldsymbol{\Gamma}_h$.

For this discussion we will find it useful to consider the different representations of the elements of the space $\boldsymbol{\Gamma}_h$ as vectors in Euclidean space, equipped with the ordinary Euclidean inner product. Let d denote the dimension of $\boldsymbol{\Gamma}_h$. If $\boldsymbol{\zeta} \in \boldsymbol{\Gamma}_h$ we let $\Pi_h \boldsymbol{\zeta} \in \mathbb{R}^d$ be the vector of coefficients in the expansion of $\boldsymbol{\zeta}$ in terms of a given local basis $\{\boldsymbol{\eta}_j\}$ of $\boldsymbol{\Gamma}_h$, i.e.,

$$\boldsymbol{\zeta} = \sum_{j=1}^{d} (\Pi_h \boldsymbol{\zeta})_j \boldsymbol{\eta}_j.$$

Furthermore, we let $\Upsilon_h : \boldsymbol{\Gamma}_h \to \mathbb{R}^d$ be defined by the inner products with the basis functions, i.e., $(\Upsilon_h \boldsymbol{\zeta})_j = (\boldsymbol{\zeta}, \boldsymbol{\eta}_j)$. By construction we have that these maps preserve inner products in the sense that

$$(\boldsymbol{\zeta}, \boldsymbol{\eta}) = (\Pi_h \boldsymbol{\zeta}) \cdot (\Upsilon_h \boldsymbol{\eta}),$$

and as a consequence,

$$\Upsilon_h^* = \Pi_h^{-1}, \tag{6.1}$$

where the asterisk is used to denote the adjoint operation. The significance of these representation operators is due to the fact that in the finite element method, the coefficient operator $\mathcal{A}_{t,h} : X_h \to X_h$ is naturally represented as a matrix, frequently referred to as the stiffness matrix, mapping the coefficients of $x \in X_h$, with respect to a given local basis, into the corresponding representation of $\mathcal{A}_{t,h} x$ in terms of inner products. More precisely,

$$\mathrm{diag}(\Upsilon_{0,h}, \Upsilon_h) \mathcal{A}_{t,h} \mathrm{diag}(\Pi_{0,h}^{-1}, \Pi_h^{-1}),$$

is a sparse matrix. Here the operators $\Pi_{0,h}$ and $\Upsilon_{0,h}$, mapping $\boldsymbol{V}_h \times W_h$ into a Euclidean space, are defined analogously to the operators Π_h and Υ_h above. Hence, the preconditioner $\mathcal{B}_{t,h}$ must have the reversed representation. In particular, this means that the desired representation for the operator $\boldsymbol{D}_{t,h}$ is the matrix $\Pi_h \boldsymbol{D}_{t,h} \Upsilon_h^{-1}$.

Let us recall that the operators $\boldsymbol{D}_{t,h} : \boldsymbol{\Gamma}_h \to \boldsymbol{\Gamma}_h$, given by (5.10), are defined from operators $\Phi_{t,h} : Q_h \to Q_h$. These operators are supposed to be preconditioners for discrete approximations of the differential operators $I - t^2\Delta$ defined on Q_h. In particular, when $t = 0$ the operator $\Phi_{0,h}$ is required to be spectrally equivalent to the identity operator on Q_h. Since the operators $\Phi_{t,h}$ correspond to preconditioners for discrete elliptic operators, it is reasonable to assume that it is inexpensive to evaluate the operators $\tilde{\Pi}_h \Phi_{t,h} \tilde{\Upsilon}_h^{-1}$, i.e., the cost is proportional to the dimension of Q_h. Hence, the desired representation of $\boldsymbol{D}_{t,h}$ should utilize this representation of $\Phi_{t,h}$. Here, the representation operators $\tilde{\Pi}_h, \tilde{\Upsilon}_h$, mapping Q_h into Euclidean space are defined from expansions and inner products with respect to a given local basis, in analogy to the operators Π_h, Υ_h on $\boldsymbol{\Gamma}_h$ defined above.

Consider first the method introduced by Arnold and Falk [1]. For this method $\boldsymbol{\Gamma}_h$ is the space of piecewise constants and Q_h is the space of continuous piecewise linear functions. Hence $\mathbf{curl}_h$ is the ordinary curl operator, but the adjoint operator rot_h is defined by (3.5). Therefore, the matrix $\Pi_h \, \mathbf{curl}_h \, \tilde{\Pi}_h^{-1}$ is local. Furthermore, by (3.5) and the property (6.1) it follows that $\tilde{\Upsilon}_h \, \mathrm{rot}_h \, \Upsilon_h^{-1}$ is the transpose matrix. Also, since $\boldsymbol{\Gamma}_h$ is the space of vector piecewise constants, the matrix $\Pi_h \Upsilon_h^{-1}$ is diagonal.

We can now evaluate $\Pi_h \boldsymbol{D}_{t,h} \Upsilon_h^{-1}$ as a composition of local operators by using the identity

$$\Pi_h \boldsymbol{D}_{t,h} \Upsilon_h^{-1} = \Pi_h \Upsilon_h^{-1} + (1 - t^2)(\Pi_h \, \mathbf{curl}_h \, \tilde{\Pi}_h^{-1})(\tilde{\Pi}_h \Phi_{t,h} \tilde{\Upsilon}_h^{-1})(\tilde{\Upsilon}_h \, \mathrm{rot}_h \, \Upsilon_h^{-1}).$$

We should remark here that even if $t = 0$, we need to replace the identity operator on Q_h by a suitable preconditioner $\Phi_{0,h}$. The reason for this is that the inverse mass matrix $\tilde{\Pi}_h \tilde{\Upsilon}_h^{-1}$ is not local for the continuous piecewise linear space Q_h. A suitable operator $\Phi_{0,h} : Q_h \to Q_h$ is in this case given as the sum of local projections, i.e.

$$\Phi_{0,h} p = \sum_j \frac{(p, q_j)}{\|q_j\|_0^2} q_j. \tag{6.2}$$

Here $\{q_j\}$ is the standard local basis on Q_h. This operator is, under suitable weak uniformity assumptions on the triangulations, spectrally equivalent to the identity operator; cf. [23]. Furthermore, $\tilde{\Pi}_h \Phi_{0,h} \tilde{\Upsilon}_h^{-1}$ is a diagonal matrix.

Consider next the case where $\boldsymbol{\Gamma}_h$ is a conforming subspace of $\mathring{\boldsymbol{H}}(\mathrm{rot})$, such as in the methods discussed in [9] and [12]. In this case rot_h is the ordinary rotation operator and $\mathrm{rot}_h(\boldsymbol{\Gamma}_h) \subset Q_h$. Again, the problem in this case is that the inverse mass matrix $\Pi_h \Upsilon_h^{-1}$ is not local. Writing $\Pi_h \boldsymbol{D}_{t,h} \Upsilon_h^{-1}$ in the form

$$\begin{aligned} \Pi_h \boldsymbol{D}_{t,h} \Upsilon_h^{-1} &= \Pi_h \Upsilon_h^{-1} \\ &\quad + (1 - t^2)(\Pi_h \Upsilon_h^{-1})(\Upsilon_h \, \mathbf{curl}_h \, \tilde{\Pi}_h^{-1})(\tilde{\Pi}_h \Phi_{t,h} \tilde{\Upsilon}_h^{-1})(\tilde{\Upsilon}_h \, \mathrm{rot}_h \, \Pi_h^{-1})(\Pi_h \Upsilon_h^{-1}), \end{aligned}$$

we note that $\tilde{\Upsilon}_h \, \mathrm{rot}_h \, \Pi_h^{-1}$ is a local operator and therefore it's adjoint $\Upsilon_h \, \mathbf{curl}_h \, \tilde{\Pi}_h^{-1}$ is also local. Hence, the only difficulty in evaluating the operator $\Upsilon_h \boldsymbol{D}_{t,h} \Pi_h^{-1}$ is that we need to evaluate the nonlocal matrix $\Pi_h \Upsilon_h^{-1}$. In the next section, we will discuss one way of overcoming this problem.

7. A relaxation procedure. The purpose of this section is to discuss the construction of preconditioners in the case where $\mathbf{\Gamma}_h \subset \mathring{\boldsymbol{H}}(\mathrm{rot})$. We recall that in this case we also have $\mathbf{grad}_h = \mathbf{grad}$, with $\mathbf{grad}(W_h) \subset \mathbf{\Gamma}_h$ and that $\boldsymbol{R}_h(\boldsymbol{V}_h) \subset \mathbf{\Gamma}_h$.

In order to avoid the evaluation of the inverse mass matrix $\Pi_h \Upsilon_h^{-1}$, which enters the expression of the preconditioner $\Pi_h \boldsymbol{D}_{t,h} \Upsilon_h^{-1}$, we will consider an extended system, posed in a larger space $\hat{\mathbf{\Gamma}}_h$. The extended system will have the same solution as the original system (3.4). However, since we have extended the space, the sequence of approximations generated by an iterative method like the minimum residual method will in general not be in the space X_h. Therefore, we refer to this approach as a relaxation procedure.

Assume that the space $\mathbf{\Gamma}_h$ is constructed from a triangulation $\mathcal{T}_h$. On each interior edge let s be a chosen unit tangent vector. Since $\mathbf{\Gamma}_h \subset \mathring{\boldsymbol{H}}(\mathrm{rot})$, $\boldsymbol{\eta} \cdot s$ is necessarily continuous on each interior edge for any $\boldsymbol{\eta} \in \mathbf{\Gamma}_h$. We let $\hat{\mathbf{\Gamma}}_h$ denote the larger, discontinuous space obtained by removing the continuity constraints on the interior edges. In particular, if $\{\boldsymbol{\eta}_j\}_{j=1}^d$ is a basis for $\mathbf{\Gamma}_h$, then we obtain a basis for $\hat{\mathbf{\Gamma}}_h$ of the form $\{\boldsymbol{\eta}_{j,T}\}_{j=1,T\in\mathcal{T}_j}^d$, where $\mathcal{T}_j$ denotes the set of triangles in the support of $\boldsymbol{\eta}_j$ and $\boldsymbol{\eta}_{j,T}$ denotes the restriction of $\boldsymbol{\eta}_j$ to T. Let $\hat{d}$ denote the dimension of $\hat{\mathbf{\Gamma}}_h$.

Let $\hat{\Pi}_h$ and $\hat{\Upsilon}_h$ be the representation operators, mapping $\hat{\mathbf{\Gamma}}_h$ into Euclidean space, which are the obvious extensions of the operators Π_h and Υ_h introduced above. Since $\hat{\mathbf{\Gamma}}_h$ is discontinuous, the inverse of the mass matrix, $\hat{\Pi}_h \hat{\Upsilon}_h^{-1}$ is a local, block diagonal operator. Let $\hat{X}_h = \boldsymbol{V}_h \times W_h \times \hat{\mathbf{\Gamma}}_h$ and define an L^2-symmetric operator $\hat{\mathcal{A}}_{t,h} : \hat{X}_h \to \hat{X}_h$ as the coefficient operator of the system (3.1)–(3.3), but where we use the space $\hat{\mathbf{\Gamma}}_h$ instead of $\mathbf{\Gamma}_h$. If $\boldsymbol{P}_h : \hat{\mathbf{\Gamma}}_h \to \mathbf{\Gamma}_h$ is the L^2-projection, then the operator $\hat{\mathcal{A}}_{t,h}$ can be alternatively written as

$$\hat{\mathcal{A}}_{t,h} = \mathcal{A}_{t,h}\, \mathcal{P}_h - t^2(\mathcal{I} - \mathcal{P}_h),$$

where $\mathcal{P}_h = \mathrm{diag}(\boldsymbol{I}, I, \boldsymbol{P}_h)$ and where $\mathcal{I}$ denotes the identity operator on $\hat{X}_h$. Hence, the operator $\hat{\mathcal{A}}_{t,h}$ is block diagonal with respect to the decomposition

$$\hat{X}_h = X_h \oplus X_h^{\perp}, \tag{7.1}$$

where $X_h^{\perp}$ is the orthogonal complement of X_h in $\hat{X}_h$ with respect to the L^2 inner product. However, the operator $\hat{\mathcal{A}}_{t,h}$ will be singular when $t = 0$. In order to avoid this difficulty, we will introduce a perturbation of this operator.

Define an operator $\boldsymbol{J}_h : \hat{\mathbf{\Gamma}}_h \to \mathbf{\Gamma}_h$ by averaging the coefficients, i.e.,

$$(\Pi_h(\boldsymbol{J}_h\boldsymbol{\eta}))_j = \frac{1}{|\mathcal{T}_j|} \sum_{T\in\mathcal{T}_j} (\hat{\Pi}_h\boldsymbol{\eta})_{j,T},$$

where $|\mathcal{T}_j|$ denotes the number of triangles in $\mathcal{T}_j$. Hence,

$$\boldsymbol{J}_h\boldsymbol{\eta} = \boldsymbol{\eta} \quad \text{for all } \boldsymbol{\eta} \in \mathbf{\Gamma}_h.$$

We also observe that the operator $\Pi_h \boldsymbol{J}_h \hat{\Pi}_h^{-1} : \mathbb{R}^{\hat{d}} \to \mathbb{R}^d$ is local. Furthermore, if the triangulation $\mathcal{T}_h$ is quasiuniform, then the operator $\boldsymbol{J}_h$ is L^2-bounded, i.e., there is a constant c, independent of h, such that

$$\|\boldsymbol{J}_h \boldsymbol{\eta}\|_0 \leq c\|\boldsymbol{\eta}\|_0 \quad \text{for all } \boldsymbol{\eta} \in \hat{\boldsymbol{\Gamma}}_h. \tag{7.2}$$

This inequality will be assumed throughout this section.

Let $\boldsymbol{\Gamma}_h^{\perp}$ be the orthogonal complement of $\boldsymbol{\Gamma}_h$ in $\hat{\boldsymbol{\Gamma}}_h$ with respect to the L^2 inner product. The norms $\|\cdot\|_0$ and $\|(\boldsymbol{I}-\boldsymbol{J}_h)\cdot\|_0$ are equivalent on $\boldsymbol{\Gamma}_h^{\perp}$, i.e., there is a constant c, independent of h, such that

$$\|\boldsymbol{\eta}\|_0 \leq \|(\boldsymbol{I}-\boldsymbol{J}_h)\boldsymbol{\eta}\|_0 \leq c\|\boldsymbol{\eta}\|_0 \quad \text{for all } \boldsymbol{\eta} \in \boldsymbol{\Gamma}_h^{\perp}.$$

The left inequality here follows since $(\boldsymbol{\eta}, \boldsymbol{J}_h\boldsymbol{\eta}) = 0$ for any $\boldsymbol{\eta} \in \boldsymbol{\Gamma}_h^{\perp}$, while the right inequality follows from (7.2). Since $(\boldsymbol{I}-\boldsymbol{J}_h)(\boldsymbol{I}-\boldsymbol{P}_h) = \boldsymbol{I}-\boldsymbol{J}_h$ this can be equivalently written as

$$\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\eta}\|_0 \leq \|(\boldsymbol{I}-\boldsymbol{J}_h)\boldsymbol{\eta}\|_0 \leq c\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\eta}\|_0 \quad \text{for all } \boldsymbol{\eta} \in \hat{\boldsymbol{\Gamma}}_h. \tag{7.3}$$

Observe that the operator $(\boldsymbol{I}-\boldsymbol{J}_h)^*(\boldsymbol{I}-\boldsymbol{J}_h)$, where $(\boldsymbol{I}-\boldsymbol{J}_h)^* : \hat{\boldsymbol{\Gamma}}_h \to \hat{\boldsymbol{\Gamma}}_h$ is the L^2-adjoint of $\boldsymbol{I}-\boldsymbol{J}_h : \hat{\boldsymbol{\Gamma}}_h \to \hat{\boldsymbol{\Gamma}}_h$, maps $\hat{\boldsymbol{\Gamma}}_h$ into $\boldsymbol{\Gamma}_h^{\perp}$. Hence, if we let $\mathcal{K}_h : \hat{X}_h \to \hat{X}_h$ be given by

$$\mathcal{K}_h = -h^2 \operatorname{diag}(0, 0, (\boldsymbol{I}-\boldsymbol{J}_h)^*(\boldsymbol{I}-\boldsymbol{J}_h))$$

then $\mathcal{K}_h$ is L^2-symmetric and maps $\hat{X}_h$ into $(X_h^{\perp})$. Instead of the system (3.4), consider the extended system

$$(\hat{\mathcal{A}}_{t,h} + \mathcal{K}_h)\begin{pmatrix}\boldsymbol{\phi}_h\\ \omega_h\\ \boldsymbol{\zeta}_h\end{pmatrix} = \begin{pmatrix}\boldsymbol{f}_h\\ g_h\\ \boldsymbol{j}_h\end{pmatrix}. \tag{7.4}$$

It follows directly from the block diagonal structure of this system, with respect to the decomposition (7.1), that it is nonsingular. Furthermore, if $\boldsymbol{j}_h \in \boldsymbol{\Gamma}_h$, then $\boldsymbol{\zeta}_h \in \boldsymbol{\Gamma}_h$ and hence, for such data, the solution of (7.4) is also a solution of (3.4).

We now propose to solve the symmetric system (7.4) by the minimum residual method (or a similar method). The coefficient operator $\hat{\mathcal{A}}_{t,h}+\mathcal{K}_h$ of the system (7.4) is represented by the matrix

$$\operatorname{diag}(\Upsilon_{0,h}, \hat{\Upsilon}_h)(\hat{\mathcal{A}}_{t,h} + \mathcal{K}_h)\operatorname{diag}(\Pi_{0,h}^{-1}, \hat{\Pi}_h^{-1}),$$

Hence, we need to construct a suitable preconditioner $\hat{\mathcal{B}}_{t,h}$ for $\hat{\mathcal{A}}_{t,h} + \mathcal{K}_h$ such that the operator

$$\operatorname{diag}(\Pi_{0,h}, \hat{\Pi}_h)\hat{\mathcal{B}}_{t,h}\operatorname{diag}(\Upsilon_{0,h}^{-1}, \hat{\Upsilon}_h^{-1}),$$

can be effectively evaluated. As above, the desired mapping property for the preconditioner will be derived from the corresponding mapping property of the coefficient operator.

In order to state the mapping property of the operator $\hat{\mathcal{A}}_{t,h}+\mathcal{K}_h$ more precisely, define in the same way as above, a t and h dependent norm on $\hat{X}_{t,h}=\hat{X}_h$ given by

$$\|(\boldsymbol{\phi},\omega,\boldsymbol{\zeta})\|^2_{\hat{X}_{t,h}} = \|(\boldsymbol{\phi},\omega,\boldsymbol{P}_h\boldsymbol{\zeta})\|^2_{X_{t,h}} + (t^2+h^2)\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\zeta}\|_0^2.$$

Furthermore, the corresponding dual space $\hat{X}^*_{t,h}$, also equal to $\hat{X}_h$ as a set, carries the dual norm:

$$\|(\boldsymbol{\phi},\omega,\boldsymbol{\zeta})\|^2_{\hat{X}^*_{t,h}} = \|(\boldsymbol{\phi},\omega,\boldsymbol{P}_h\boldsymbol{\zeta})\|^2_{X^*_{t,h}} + \frac{1}{t^2+h^2}\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\zeta}\|_0^2.$$

It follows directly from the mapping properties of $\mathcal{A}_{t,h}$ and the block diagonal structure of $\hat{\mathcal{A}}_{t,h}+\mathcal{K}_h$ with respect to the decomposition (7.1) that the operator norms

$$\|\hat{\mathcal{A}}_{t,h}+\mathcal{K}_h\|_{\mathcal{L}(\hat{X}_{t,h},\hat{X}^*_{t,h})} \qquad \text{and} \qquad \|(\hat{\mathcal{A}}_{t,h}+\mathcal{K}_h)^{-1}\|_{\mathcal{L}(\hat{X}^*_{t,h},\hat{X}_{t,h})}$$

are bounded independently of t and h, and this determines the desired mapping properties of a possible preconditioner $\hat{\mathcal{B}}_{t,h}=\operatorname{diag}(\boldsymbol{L}_h, M_h, \hat{\boldsymbol{N}}_{t,h})$.

As above, the desired properties of the operators $\boldsymbol{L}_h$ and M_h are given by (5.4) and (5.5). Hence, we only need to consider the operator $\hat{\boldsymbol{N}}_{t,h}:\hat{\boldsymbol{\Gamma}}_h\to\hat{\boldsymbol{\Gamma}}_h$. From the properties of $\hat{\mathcal{A}}_{t,h}+\mathcal{K}_h$ it follows that $\hat{\boldsymbol{N}}_{t,h}$ is required to have operator norms

$$\|\hat{\boldsymbol{N}}_{t,h}\|_{\mathcal{L}(\hat{\boldsymbol{\Gamma}}^*_{t,h},\hat{\boldsymbol{\Gamma}}_{t,h})} \qquad \text{and} \qquad \|\hat{\boldsymbol{N}}^{-1}_{t,h}\|_{\mathcal{L}(\hat{\boldsymbol{\Gamma}}_{t,h},\hat{\boldsymbol{\Gamma}}^*_{t,h})} \tag{7.5}$$

independent of t and h. Here the norms $\|\cdot\|_{\hat{\boldsymbol{\Gamma}}_{t,h}}$ and $\|\cdot\|_{\hat{\boldsymbol{\Gamma}}^*_{t,h}}$ are defined by restricting the norms $\|\cdot\|_{\hat{X}_{t,h}}$ and $\|\cdot\|_{\hat{X}^*_{t,h}}$ to $\hat{\boldsymbol{\Gamma}}_h$, i.e.,

$$\|\boldsymbol{\zeta}\|^2_{\hat{\boldsymbol{\Gamma}}_{t,h}} = \|\boldsymbol{P}_h\boldsymbol{\zeta}\|^2_{\boldsymbol{\Gamma}_{t,h}} + (t^2+h^2)\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\zeta}\|_0^2$$

and

$$\|\boldsymbol{\zeta}\|^2_{\hat{\boldsymbol{\Gamma}}^*_{t,h}} = \|\boldsymbol{P}_h\boldsymbol{\zeta}\|^2_{\boldsymbol{\Gamma}^*_{t,h}} + \frac{1}{t^2+h^2}\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\zeta}\|_0^2,$$

where $\|\cdot\|_{\boldsymbol{\Gamma}_{t,h}} = \|\cdot\|_{\boldsymbol{H}_h^{-1}(\mathrm{div}_h)\cap t\cdot\boldsymbol{L}^2}$ and $\|\cdot\|_{\boldsymbol{\Gamma}^*_{t,h}} = \|\cdot\|_{\mathring{\boldsymbol{H}}_h(\mathrm{rot}_h)+t^{-1}\cdot\boldsymbol{L}^2}$.

In order to construct a suitable preconditioner $\hat{\boldsymbol{N}}_{t,h}$ satisfying (7.5), we will utilize the operators $\boldsymbol{D}_{t,h}:\boldsymbol{\Gamma}_h\to\boldsymbol{\Gamma}_h$ given in (5.10). By Theorem 5.2, the operator norms

$$\|\boldsymbol{D}_{t,h}\|_{\mathcal{L}(\boldsymbol{\Gamma}^*_{t,h},\boldsymbol{\Gamma}_{t,h})} \qquad \text{and} \qquad \|\boldsymbol{D}^{-1}_{t,h}\|_{\mathcal{L}(\boldsymbol{\Gamma}_{t,h},\boldsymbol{\Gamma}^*_{t,h})} \tag{7.6}$$

are bounded independently of t and h.

Consider the operator

$$\boldsymbol{D}_{t,h}\boldsymbol{P}_h + \frac{1}{t^2+h^2}(\boldsymbol{I}-\boldsymbol{P}_h) \tag{7.7}$$

on $\hat{\boldsymbol{\Gamma}}_h$. It follows directly from (7.6) and the definitions of the norms $\|\cdot\|_{\hat{\boldsymbol{\Gamma}}_{t,h}}$ and $\|\cdot\|_{\hat{\boldsymbol{\Gamma}}^*_{t,h}}$ that this operator has the mapping property (7.5) required for $\hat{\boldsymbol{N}}_{t,h}$. However, this operator will not be computationally effective since the appearance of the L^2-projection makes it necessary to evaluate the inverse mass matrix $\Pi_h\Upsilon_h^{-1}$.

Instead of the operator (7.7) we will therefore define $\hat{\boldsymbol{N}}_{t,h}:\hat{\boldsymbol{\Gamma}}_h \to \hat{\boldsymbol{\Gamma}}_h$ by

$$\hat{\boldsymbol{N}}_{t,h} = \boldsymbol{J}_h^*\boldsymbol{D}_{t,h}\boldsymbol{J}_h + \frac{1}{t^2+h^2}(\boldsymbol{I}-\boldsymbol{J}_h)^*(\boldsymbol{I}-\boldsymbol{J}_h). \tag{7.8}$$

Since the matrices $\Pi_h\boldsymbol{J}_h\hat{\Pi}_h^{-1}$, $\hat{\Upsilon}_h\boldsymbol{J}_h^*\Upsilon_h^{-1}$, and $\hat{\Pi}_h\hat{\Upsilon}_h^{-1}$ are local, the operator $\hat{\Pi}_h\hat{\boldsymbol{N}}_{t,h}\hat{\Upsilon}_h^{-1}$ can be expressed as the product of local matrices according to the formula

$$\begin{aligned}\hat{\Pi}_h\hat{\boldsymbol{N}}_{t,h}\hat{\Upsilon}_h^{-1} = {} & (\hat{\Pi}_h\hat{\Upsilon}_h^{-1})(\hat{\Upsilon}_h\boldsymbol{J}_h^*\Upsilon_h^{-1})\big[(\Upsilon_h\Pi_h^{-1}) \\ & +(1-t^2)(\Upsilon_h\,\mathbf{curl}_h\,\tilde{\Pi}_h^{-1})(\tilde{\Pi}_h\Phi_{t,h}\tilde{\Upsilon}_h^{-1})(\tilde{\Upsilon}_h\,\mathrm{rot}_h\,\Pi_h^{-1})\big](\Pi_h\boldsymbol{J}_h\hat{\Pi}_h^{-1})(\hat{\Pi}_h\hat{\Upsilon}_h^{-1}) \\ & +\frac{1}{t^2+h^2}(\hat{\Pi}_h\hat{\Upsilon}_h^{-1})(\hat{\Upsilon}_h[\boldsymbol{I}-\boldsymbol{J}_h]^*\Upsilon_h^{-1})(\Upsilon_h\Pi_h^{-1})(\Pi_h[\boldsymbol{I}-\boldsymbol{J}_h]\hat{\Pi}_h^{-1})(\hat{\Pi}_h\hat{\Upsilon}_h^{-1}),\end{aligned}$$

and hence can be effectively computed.

We also need to show that the operator $\hat{\boldsymbol{N}}_{t,h}$ is spectrally equivalent to the operator given by (7.7). As was done in Lemma 5.3, we shall assume that the space $\boldsymbol{\Gamma}_h$ admits an inverse property of the form

$$\|\operatorname{rot}\boldsymbol{\eta}\|_0 \le K_1 h^{-1}\|\boldsymbol{\eta}\|_0, \quad \text{for all } \boldsymbol{\eta}\in\boldsymbol{\Gamma}_h, \tag{7.9}$$

where the constant K_1 is independent of h.

The following result implies that the operator $\hat{\boldsymbol{N}}_{t,h}$ satisfies the mapping property (7.5).

Lemma 7.1. *Assume that the property* (7.9) *holds and that the operator* $\boldsymbol{D}_{t,h}:\boldsymbol{\Gamma}_h\to\boldsymbol{\Gamma}_h$ *is defined by* (5.10). *Then the operator* $\hat{\boldsymbol{N}}_{t,h}$, *defined by* (7.8), *is spectrally equivalent, uniformly in* t *and* h, *to the operator given by* (7.7).

Proof. We first establish that the spectral radius, $\rho(\boldsymbol{D}_{t,h})$, of $\boldsymbol{D}_{t,h}$ satisfies the bound

$$\rho(\boldsymbol{D}_{t,h}) \le c(t^2+h^2)^{-1}, \tag{7.10}$$

where c is independent of t and h. Since from the proof of Theorem 5.2, the operators $\boldsymbol{D}_{t,h}$ and $\boldsymbol{\Lambda}_{1,h}\boldsymbol{\Lambda}_{t,h}^{-1}$ are spectrally equivalent, this bound would follow if a corresponding property holds for the operator $\mathbf{\Lambda}_{1,h}\mathbf{\Lambda}_{t,h}^{-1}$. Note, however, that if $\boldsymbol{\eta}$ is an eigenfunction of $\mathbf{\Lambda}$ with eigenvalue μ, then $\boldsymbol{\eta}$ is also an eigenfunction of $\mathbf{\Lambda}_{t,h}$ and $\mathbf{\Lambda}_{1,h}\mathbf{\Lambda}_{t,h}^{-1}$ with eigenvalues $1-t^2+t^2\mu$ and $\mu/(1-t^2+t^2\mu) \equiv \lambda$, respectively. Hence,

$$\Lambda\boldsymbol{\eta} = \mu\boldsymbol{\eta} = \frac{\mu}{1-t^2+t^2\mu}(1-t^2+t^2\mu)\boldsymbol{\eta} = \lambda\mathbf{\Lambda}_{t,h}\boldsymbol{\eta}.$$

Taking inner products with $\boldsymbol{\eta}$, we get

$$\|\boldsymbol{\eta}\|_0^2 + \|\operatorname{rot}\boldsymbol{\eta}\|_0^2 = \lambda(\|\boldsymbol{\eta}\|_0^2 + t^2\|\operatorname{rot}\boldsymbol{\eta}\|_0^2).$$

Hence, $\lambda = f(\|\operatorname{rot}\boldsymbol{\eta}\|_0^2/\|\boldsymbol{\eta}\|_0^2)$, where $f(x) = (1+x)/(1+t^2x)$. Since $f(x)$ is an increasing function for $0 \le t \le 1$ and $\|\operatorname{rot}\boldsymbol{\eta}\|_0^2/\|\boldsymbol{\eta}\|_0^2 \le K_1^2h^{-2}$ by (7.9), we get that

$$\lambda \le \frac{1+K_1^2h^{-2}}{1+t^2K_1^2h^{-2}} \le c(t^2+h^2)^{-1}.$$

Hence,

$$\rho(\mathbf{\Lambda}_{1,h}\mathbf{\Lambda}_{t,h}^{-1}) \le c(t^2+h^2)^{-1},$$

and this implies (7.10).

To show the spectral equivalence, we observe that

$$\begin{aligned}
(\boldsymbol{J}_h^*\boldsymbol{D}_{t,h}\boldsymbol{J}_h\boldsymbol{\eta},\boldsymbol{\eta}) &+ \frac{1}{t^2+h^2}([\boldsymbol{I}-\boldsymbol{J}_h]^*[\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta},\boldsymbol{\eta}) \\
&= (\boldsymbol{D}_{t,h}\boldsymbol{J}_h\boldsymbol{\eta},\boldsymbol{J}_h\boldsymbol{\eta}) + \frac{1}{t^2+h^2}([\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta},[\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta}) \\
&= (\boldsymbol{D}_{t,h}\boldsymbol{P}_h\boldsymbol{\eta},\boldsymbol{J}_h\boldsymbol{\eta}) + (\boldsymbol{D}_{t,h}[\boldsymbol{J}_h-\boldsymbol{P}_h]\boldsymbol{\eta},\boldsymbol{J}_h\boldsymbol{\eta}) + \frac{1}{t^2+h^2}([\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta},[\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta}) \\
&\le [(\boldsymbol{D}_{t,h}\boldsymbol{P}_h\boldsymbol{\eta},\boldsymbol{P}_h\boldsymbol{\eta})^{1/2} + (\boldsymbol{D}_{t,h}[\boldsymbol{J}_h-\boldsymbol{P}_h]\boldsymbol{\eta},[\boldsymbol{J}_h-\boldsymbol{P}_h]\boldsymbol{\eta})^{1/2}](\boldsymbol{D}_{t,h}\boldsymbol{J}_h\boldsymbol{\eta},\boldsymbol{J}_h\boldsymbol{\eta})^{1/2} \\
&\quad + \frac{1}{t^2+h^2}([\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta},[\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta}).
\end{aligned}$$

Hence, using the arithmetic-geometric mean inequality, (7.10), (7.3), and the triangle inequality, we get

$$\begin{aligned}
&(\boldsymbol{J}_h^*\boldsymbol{D}_{t,h}\boldsymbol{J}_h\boldsymbol{\eta},\boldsymbol{\eta}) + \frac{1}{t^2+h^2}([\boldsymbol{I}-\boldsymbol{J}_h]^*[\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta},\boldsymbol{\eta}) \\
&\le 2(\boldsymbol{D}_{t,h}\boldsymbol{P}_h\boldsymbol{\eta},\boldsymbol{P}_h\boldsymbol{\eta}) + 2(\boldsymbol{D}_{t,h}[\boldsymbol{J}_h-\boldsymbol{P}_h]\boldsymbol{\eta},[\boldsymbol{J}_h-\boldsymbol{P}_h]\boldsymbol{\eta}) + \frac{2}{t^2+h^2}([\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta},[\boldsymbol{I}-\boldsymbol{J}_h]\boldsymbol{\eta}) \\
&\le 2(\boldsymbol{D}_{t,h}\boldsymbol{P}_h\boldsymbol{\eta},\boldsymbol{P}_h\boldsymbol{\eta}) + \frac{c}{t^2+h^2}\|(\boldsymbol{J}_h-\boldsymbol{P}_h)\boldsymbol{\eta}\|_0^2 + \frac{c}{t^2+h^2}\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\eta}\|_0^2 \\
&\le c\left[(\boldsymbol{D}_{t,h}\boldsymbol{P}_h\boldsymbol{\eta},\boldsymbol{P}_h\boldsymbol{\eta}) + \frac{1}{t^2+h^2}\|(\boldsymbol{I}-\boldsymbol{P}_h)\boldsymbol{\eta}\|_0^2\right] \\
&= c([\boldsymbol{D}_{t,h}\boldsymbol{P}_h + \frac{1}{t^2+h^2}(\boldsymbol{I}-\boldsymbol{P}_h)]\boldsymbol{\eta},\boldsymbol{\eta}).
\end{aligned}$$

The reverse inequality follows by a similar argument. □

8. Numerical examples. In this section, we shall report on numerical experiments using some of the preconditioners developed in this paper. The main purpose of these experiments is to illustrate the typical behavior of the algorithms discussed above. Therefore, we have considered only modest size problems using simple meshes. In fact, all the experiments are done in Matlab. The largest systems we consider below have approximately $250,000$ unknowns.

The material constants are chosen such that $k = 5/6$, $E = 3$, and $\nu = 1/4$. Hence, $\lambda = Ek/2(1 + \nu) = 1$.

The domain $\Omega \subset \mathbb{R}^2$ is taken to be the unit square. The triangulation of Ω is obtained by first dividing Ω into squares of size $h \times h$, and then dividing each square into two triangles using the positively sloped diagonal. Furthermore, all the computations are done with the method of Arnold and Falk described in §3 above. Hence, the space $\boldsymbol{V}_h$ consists of continuous piecewise linear functions plus cubic bubbles on each triangle, W_h is the nonconforming piecewise linear space, with continuity requirements only at the midpoint of each edge, and $\boldsymbol{\Gamma}_h$ is the space of piecewise constants. Furthermore, the auxiliary space Q_h, which will be needed in order to construct the operators $\boldsymbol{D}_{t,h}$, is the space of continuous piecewise linear functions.

As discussed above, we shall consider the preconditioned system (5.1) with a block diagonal preconditioner $\mathcal{B}_{t,h}$ of the form (5.3), i.e.,

$$\mathcal{B}_{t,h} = \operatorname{diag}(\boldsymbol{L}_h, M_h, \boldsymbol{N}_{t,h}).$$

In order to define the proper preconditioners $\boldsymbol{L}_h$ and M_h, defined on $\boldsymbol{V}_h$ and W_h respectively, we shall utilize a preconditioner Ψ_h for the discrete Laplace operator, with Dirichlet boundary conditions, defined on the corresponding conforming space, $W_h^c \subset W_h$, consisting of continuous piecewise linear functions. In our examples below, the preconditioner $\Psi_h : W_h^c \to W_h^c$ is a standard V-cycle multigrid operator, where a Gauss-Seidel operator is used as a smoother and where the coarsest level corresponds to $h = 1/2$.

In the space $\boldsymbol{V}_h$, the subspace spanned by the bubble functions is orthogonal to the space of continuous piecewise linear functions, $\boldsymbol{W}_h^c$, with respect to the Dirichlet form. Hence, the operator $\boldsymbol{L}_h$ can be constructed from a diagonal matrix, corresponding to the space spanned by the bubble functions, and by two copies of the operator Ψ_h. The operator M_h, on the nonconforming space W_h, is constructed from Ψ_h by using the auxiliary space approach described in Xu [25]. Hence, the operator M_h is of the form

$$M_h = R_h + \Psi_h P_h^c,$$

where $P_h^c : W_h \to W_h^c$ is the L^2-projection and where $R_h : W_h \to W_h$ is a smoothing operator. In the examples below, R_h is obtained from two Richardson iterations. The preconditioners $\boldsymbol{L}_h$ and M_h will be fixed throughout all the examples below.

In order to construct the third block of the preconditioner, $\boldsymbol{N}_{t,h}$, we will need operators $\boldsymbol{D}_{t,h} : \boldsymbol{\Gamma}_h \to \boldsymbol{\Gamma}_h$ of the form given by (5.10). Furthermore, the definition of these operators require other operators $\Phi_{t,h} : Q_h \to Q_h$, which are preconditioners for the discretization of the Neumann problem associated the operator $I - t^2\Delta$ with respect to the conforming piecewise linear space Q_h. The operators $\Phi_{t,h}$ will be constructed analogously to the operator Ψ_h described above, i.e., $\Phi_{t,h}$ are standard V-cycle multigrid operators. However, compared to Ψ_h, we need modifications due to the t–dependent differential operator and due to the different boundary conditions.

In the examples below, the preconditioned system (5.1) is solved either by the minimum residual method or by the conjugate gradient method applied to the normal equations. Here the normal equations are defined with respect to the inner product $(\mathcal{B}_{t,h}^{-1}\,\cdot\,,\,\cdot\,)$. Hence, both methods minimize the norm associated with the inner product (5.2) over proper Krylov spaces. We recall that the work estimate for one iteration of the conjugate gradient method applied to the normal equations corresponds roughly to two minimum residual iterations. In the examples below, we therefore compare the number of iterations for the minimum residual method (N_{MR}) with twice the number of iterations for the conjugate gradient method applied to the normal equations (N_{CGN}). The condition number of the operator $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$, $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$, which we estimate from the conjugate gradient iteration using a standard Matlab routine, will also be given.

The iterations are terminated when the error, measured in the norm associated with the inner product (5.2), is reduced by a factor less than $5 \cdot 10^{-4}$.

Example 8.1. In this example $t = 0$. The preconditioner $\mathcal{B}_{t,h}$ is obtained as explained above with $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$ given by (5.10). For each of several values of h, the preconditioned system was solved by the minimum residual method and by the conjugate gradient method applied to the normal equations. The results are given in Table 1.

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
N_{MR}	41	41	35	29	24
N_{CGN}	48	50	48	40	34
$\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$	8.17	10.7	11.1	10.6	9.62

Table 1. The case $t = 0$, $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$.

Observe that, in agreement with the theory above, the condition number $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ appears to be bounded independently of h, and hence the numbers of iterations for both methods remain bounded.

Example 8.2. According to Corollary 5.4, when t is sufficiently small relative to h, then the choice $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$ should be a good one. In order to test this, we show in Table 2 the results of taking $t = 0.01$ and $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$.

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
N_{MR}	39	35	28	40	72
N_{CGN}	48	50	48	108	360
$\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$	8.15	10.7	11.4	32.9	113

TABLE 2. The case $t = 0.01$, $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$.

As expected, we observe that, for sufficiently large values of h, the choice $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$ leads to a reasonably good preconditioner. However, when h becomes small enough, so that t is no longer small compared to h, the numbers of iterations increase rapidly. This clearly illustrates that by using this simplified preconditioner, we do not get a condition number for the preconditioned system which is bounded independently of h.

We compared the results in Table 2 above with the choice $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{t,h}$. In this case Theorem 5.2 predicts that the preconditioner is uniform with respect to h. This is clearly confirmed by the results given in Table 3.

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
N_{MR}	39	36	28	25	24
N_{CGN}	48	50	48	42	36
$\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$	8.15	10.7	11.2	11.1	9.68

TABLE 3. The case $t = 0.01$, $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{t,h}$.

EXAMPLE 8.3. We next consider the case when $t = 1$ and choose $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{1,h} = \boldsymbol{I}$. We expect this choice to be a uniform preconditioner with respect to h. The results which were obtained are given in Table 4.

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
N_{MR}	22	22	20	20	20
N_{CGN}	102	104	106	104	102
$\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$	17.5	18.4	19.0	19.0	18.9

TABLE 4. The case $t = 1$, $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{t,h} = \boldsymbol{I}$.

As expected the results appear to be uniform with respect to h. Also observe the substantial difference in the behavior of the minimum residual method and the conjugate gradient method for the normal equations in this case.

Example 8.4. If we now consider the case $t = 0.1$, then by Corollary 5.5, we expect that the preconditioner $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{1,h} = \boldsymbol{I}$ used in the previous example would still be a good one. The results of that experiment are shown in Table 5.

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
N_{MR}	78	80	80	78	78
N_{CGN}	226	214	198	190	188
$\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$	90.2	78.5	72.7	70.1	70.7

Table 5. The case $t = 0.1$, $\boldsymbol{N}_{t,h} = \boldsymbol{I}$.

These results clearly reflect the fact that when $\boldsymbol{N}_{t,h} = \boldsymbol{I}$ the condition number of $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$ is independent of h. However, compared to Example 8.3, the condition numbers $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ have been substantially increased. This illustrates the dependence of the condition number $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ on t with this choice of preconditioner.

As a comparison, we repeated the experiment, but with $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{t,h}$ instead of $\boldsymbol{N}_{t,h} = \boldsymbol{I}$. The results are given in Table 6.

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
N_{MR}	28	28	27	26	26
N_{CGN}	48	54	52	50	50
$\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$	8.64	10.8	11.1	11.2	11.1

Table 6. The case $t = 0.1$, $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{t,h}$.

We observe that the condition numbers obtained in this case are comparable to what we got in Example 8.1 when $t = 0$.

In the four examples above, we have tested different choices of preconditioners $\mathcal{B}_{t,h}$, obtained by letting $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{s,h}$ for proper values of s. According to Theorem 5.2 the choice $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{t,h}$ gives a preconditioner which is uniform with respect to the thickness parameter t and the discretization parameter h. However, the operators $\boldsymbol{D}_{t,h}$ simplify in the two extreme cases, $t = 0$ and $t = 1$, since $\boldsymbol{D}_{1,h} = \boldsymbol{I}$ and since $\boldsymbol{D}_{0,h}$ can be defined from an approximation of the identity, $\Phi_{0,h}$, of the form (6.2). Hence, in the two extreme cases we do not need to implement preconditioners $\Phi_{t,h}$ for discrete versions of the operator $I - t^2\Delta$.

When $t = 0$, the results for $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$ seem to confirm the prediction that the condition numbers $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ are independent of h. In fact, the results of Example 8.1 show that the condition numbers in this case, which corresponds to the biharmonic equation, are close to 10. Hence, we have clearly demonstrated that standard preconditioners for second order elliptic operators, together with the discrete differential operator $\boldsymbol{D}_{0,h}$, is sufficient to construct an effective preconditioner for this problem. When t is sufficiently small compared to h, we have also shown, in agreement with Corollary 5.4, that $\boldsymbol{N}_{t,h} = \boldsymbol{D}_{0,h}$ is an effective preconditoner.

Furthermore, the experiments show that for a fixed $t > 0$ and $\boldsymbol{N}_{t,h} = \boldsymbol{I}$, the condition numbers $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ appears to be bounded independently of h. However, they grow with decreasing values of t. Finally, the experiments also confirm that in order to obtain a preconditioner $\mathcal{B}_{t,h}$ such that the condition number $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ is bounded independently of both the parameters t and h, we need to implement the full operator $\boldsymbol{D}_{t,h}$, given by (5.10), for the proper value of t.

Acknowledgements. The authors would like to thank Torgeir Rusten for his help with the computations reported in this section.

REFERENCES

[1] D. N. Arnold and R. S. Falk, *A uniformly accurate finite element method for the Reissner–Mindlin plate*, SIAM J. Numer. Anal., 26 (1989), pp. 1276–1290.

[2] J. Bergh and J. Löfstrom, *Interpolation spaces, an introduction*, Springer Verlag (1976).

[3] D. Braess and C. Blömer, *A multigrid method for a parameter dependent problem in solid mechanics*, Numer. Math., 57 (1990), pp. 747–762.

[4] J. H. Bramble and J. E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp., 50 (1988), pp. 1–17.

[5] J. H. Bramble and J. E. Pasciak, *Iterative techniques for time dependent Stokes problem*, to appear in Comput. Math. Appl..

[6] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev, *Analysis of inexact Uzawa algorithm for saddle point problems*, to appear in SIAM J. Numer. Anal..

[7] J. H. Bramble, J. E. Pasciak and J. Xu, *The analysis of of multigrid algorithms with nonnested spaces and noninherited quadratic forms*, Math. Comp., 56 (1991), pp. 1–34.

[8] S. C. Brenner, *Multigrid methods for parameter dependent problems*, to appear in Math. Modelling Numer. Anal. (1996).

[9] F. Brezzi, M. Fortin, and R. Stenberg, *Error analysis of mixed-interpolated elements for Reissner–Mindlin plates*, Math. Models and Methods in Applied Sciences, 1 (1991), pp. 125–151.

[10] L. C. Cowsar, *Dual variable Schwarz methods for mixed finite elements*, Report TR93–09, Rice University, Houston (1993).

[11] L. C. Cowsar, J. Mandel and M. F. Wheeler, *Balancing domain decomposition for mixed finite elements*, Math. Comp., 64 (1995), pp. 989–1015.

[12] R. Durán and E. Liberman, *On mixed finite element methods for the Reissner–Mindlin plate model*, Math. Comp., 58 (1992), pp. 561–573.

[13] H. C. Elman and G. Golub, *Inexact and preconditioned Uzawa algorithms for saddle point problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1645–1661.

[14] W. Hackbusch, *Iterative solution of large sparse systems of equations*, Springer Verlag (1994).

[15] Z. Huang, *A multi-grid algorithm for mixed problems with penalty*, Numer. Math., 57 (1990), pp. 227–247.

[16] A. Klawonn, *An optimal preconditioner for a class of saddle point problems with a penalty term*, Preprint (1994).

[17] P. Peisker, *A multigrid method for Reissner–Mindlin plates*, Numer. Math., 59 (1991), pp. 511–528.

[18] T. Rusten, P. S. Vassilevski and R. Winther, *Interior penalty preconditioners for mixed finite element approximations of elliptic problems*, Math. Comp., 65 (1996), pp. 447–466.

[19] T. Rusten and R. Winther, *A preconditioned iterative method for saddle point problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.

[20] T. Rusten and R. Winther, *Substructure preconditioners for ellptic saddle point problems*, Math. Comp., 60 (1993), pp. 23–48.

[21] D. Silvester and A. Wathen, *Fast iterative solution of stabilised Stokes systems, Part II: Using general block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367.

[22] P. S. Vassilevski and J. Wang, *An application of the abstract multilevel theory to nonconforming finite element methods*, SIAM J. Numer. Anal., 32 (1995), pp. 235–248.

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

[24] A. Wathen and D. Silvester, *Fast iterative solution of stabilised Stokes systems, Part I: Using simple diagonal preconditioners*, SIAM J. Numer. Anal., 30 (1993), pp. 630–649.

[25] J. Xu, *The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids*, to appear in Computing, 56 (1996).