MATHEMATICS OF COMPUTATION
Volume 66, Number 219, July 1997, Pages 957–984
S 0025-5718(97)00826-0

PRECONDITIONING IN $\boldsymbol{H}(\mathrm{div})$ AND APPLICATIONS

DOUGLAS N. ARNOLD, RICHARD S. FALK, AND R. WINTHER

Dedicated to Professor Ivo Babuška on the occasion of his seventieth birthday.

ABSTRACT. We consider the solution of the system of linear algebraic equations which arises from the finite element discretization of boundary value problems associated to the differential operator $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$. The natural setting for such problems is in the Hilbert space $\boldsymbol{H}(\mathrm{div})$ and the variational formulation is based on the inner product in $\boldsymbol{H}(\mathrm{div})$. We show how to construct preconditioners for these equations using both domain decomposition and multigrid techniques. These preconditioners are shown to be spectrally equivalent to the inverse of the operator. As a consequence, they may be used to precondition iterative methods so that any given error reduction may be achieved in a finite number of iterations, with the number independent of the mesh discretization. We describe applications of these results to the efficient solution of mixed and least squares finite element approximations of elliptic boundary value problems.

1. INTRODUCTION

The Hilbert space $\boldsymbol{H}(\mathrm{div})$ consists of square-integrable vectorfields on a domain Ω with square-integrable divergence. This space arises naturally in the variational formulation of a variety of systems of partial differential equations. The inner product in $\boldsymbol{H}(\mathrm{div})$ is given by

$$\Lambda(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}) + (\mathrm{div}\,\boldsymbol{u}, \mathrm{div}\,\boldsymbol{v}),$$

where $(\,\cdot\,,\,\cdot\,)$ is used to denote the inner product in L^2. Associated to the inner product Λ is a linear operator $\boldsymbol{\Lambda}$ mapping $\boldsymbol{H}(\mathrm{div})$ isometrically onto its dual space, given by the equations

$$(\boldsymbol{\Lambda}\boldsymbol{u}, \boldsymbol{v}) = \Lambda(\boldsymbol{u}, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}).$$

Just as the corresponding operator for the inner product in the Sobolev space H^1 may be considered as a realization of the differential operator $I - \Delta$ together with a homogeneous natural boundary condition, $\boldsymbol{\Lambda}$ may be thought of as a realization of the operator $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$ with an appropriate boundary condition. More precisely,

Received by the editor March 19, 1996 and, in revised form, April 19, 1996.
1991 *Mathematics Subject Classification.* Primary 65N55, 65N30 .
Key words and phrases. Preconditioner, mixed method, least squares, finite element, multigrid, domain decomposition.
The first author was supported by 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.

if $\boldsymbol{f} \in \boldsymbol{L}^2$, then the operator equation $\boldsymbol{\Lambda u} = \boldsymbol{f}$ is equivalent to the differential equation

$$\boldsymbol{u} - \mathbf{grad}\,\mathrm{div}\,\boldsymbol{u} = \boldsymbol{f} \text{ in } \Omega$$

together with the natural boundary condition

$$\mathrm{div}\,\boldsymbol{u} = 0 \text{ on } \partial\Omega.$$

Note that this is *not* an elliptic boundary value problem. When restricted to the subspace of gradient fields in $\boldsymbol{H}(\mathrm{div})$, $\boldsymbol{\Lambda}$ coincides with the second order elliptic operator $\boldsymbol{I} - \boldsymbol{\Delta}$, while when restricted to the subspace of curl fields, $\boldsymbol{\Lambda}$ coincides with the identity.

Given a finite element subspace $\boldsymbol{V}_h$ of $\boldsymbol{H}(\mathrm{div})$, we determine a positive-definite symmetric operator $\boldsymbol{\Lambda}_h : \boldsymbol{V}_h \to \boldsymbol{V}_h$ by

$$(\boldsymbol{\Lambda}_h \boldsymbol{u}, \boldsymbol{v}) = \Lambda(\boldsymbol{u}, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in \boldsymbol{V}_h. \tag{1.1}$$

Then for any $\boldsymbol{f} \in \boldsymbol{V}_h$, the equation

$$\boldsymbol{\Lambda}_h \boldsymbol{u} = \boldsymbol{f} \tag{1.2}$$

admits a unique solution $\boldsymbol{u} \in \boldsymbol{V}_h$. Once a basis for $\boldsymbol{V}_h$ has been selected, this equation may be realized on a computer as a matrix equation. Our goal in this paper is the specification and theoretical justification of fast algorithms for solving this matrix equation. Specifically we shall show how either domain decomposition or multigrid techniques can be used to efficiently construct an L^2-symmetric operator $\boldsymbol{\Theta}_h : \boldsymbol{V}_h \to \boldsymbol{V}_h$ which is spectrally equivalent to the inverse of $\boldsymbol{\Lambda}_h$, i.e., such that the spectrum of $\boldsymbol{\Theta}_h \boldsymbol{\Lambda}_h$ is bounded above and below by positive constants independent of the mesh discretization parameter h. It follows that the operator equation (1.2) can be solved efficiently by the conjugate gradient or other iterative methods using $\boldsymbol{\Theta}_h$ as a preconditioner. More precisely, the number of iterations needed to achieve a given order of accuracy will depend only on the spectral bounds, and so will not increase as the mesh is refined.

Our interest in the efficient solution of (1.2) is motivated by its applications to numerous problems of practical import. As a very simple example, consider the computation of $\boldsymbol{u} = \mathbf{grad}\, p$ where p is determined by the Dirichlet problem $-\Delta p + p = g$ in Ω, $p = 0$ on $\partial\Omega$. Then $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div})$ satisfies $\Lambda(\boldsymbol{u}, \boldsymbol{v}) = -(g, \mathrm{div}\,\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div})$. Approximating $\boldsymbol{u}$ by $\boldsymbol{u}_h \in \boldsymbol{V}_h$ and restricting $\boldsymbol{v}$ to the same space gives (1.2) where $(\boldsymbol{f}, \boldsymbol{v}) = -(g, \mathrm{div}\,\boldsymbol{v})$. (We also remark that p can be computed from $\boldsymbol{u}$ as $g + \mathrm{div}\,\boldsymbol{u}$.)

In § 7 we will consider some more significant applications of (1.2). One such application is the solution of the linear algebraic system arising from a mixed finite element discretization of a scalar second order elliptic problem. Mixed finite element methods for such problems have been widely studied and applied, but the solution of the linear algebraic systems they engender is not straightforward. As we shall show in § 7, this indefinite algebraic system has the same mapping properties as the block diagonal system whose blocks are $\boldsymbol{\Lambda}_h$ and the identity. (More precisely, the spectrum of the product of the inverse of this block diagonal operator and the indefinite operator arising from the mixed system is bounded above and below and bounded away from zero uniformly in h.) It then follows easily that if the system is preconditioned with a block diagonal preconditioner with blocks $\boldsymbol{\Theta}_h$ and the identity, then appropriate preconditioned iterative methods converge, with the

number of iterations needed to achieve a given error reduction independent of the mesh size.

Another direct application of our results which will be discussed in § 7 is the preconditioning of first order least-squares formulations of second order elliptic problems. The simplest such formulation, as discussed for example, in [25] and [13], leads to a system which has the same mapping properties as the block diagonal operator where the blocks may now be taken to be $\boldsymbol{\Lambda}_h$ and a discrete Laplacian. Hence a simple combination of our preconditioner with a standard preconditioner for the Laplacian will result in an efficient numerical method.

We mention several other applications of our results which will not be discussed below. One is to the implementation of the sequential regularization method for the nonstationary incompressible Navier–Stokes system, introduced in [23]. As discussed in § 1 of [23], the SRM iterative method requires the solution of an equation of the form (1.2) at each timestep. There are also connections between our results and iterative solvers for the Reissner–Mindlin plate [1] and with the construction of fictitious domain preconditioners for the mixed finite element method [30]. Our construction and analysis of the preconditioner $\boldsymbol{\Theta}_h$ is guided by the modern theory of multilevel and domain decomposition methods, as presented, for example, in [3] and [34]. However the operator $\boldsymbol{\Lambda}$ lacks a number of properties possessed by standard elliptic operators, and this necessitates a number of modifications to the theory. For example, in multigrid analysis it is often required that the eigenfunctions corresponding to the lowest eigenvalues of the operator can be well represented on a coarse mesh. This property is not true for $\boldsymbol{\Lambda}$ (since the curl of a highly oscillatory function is an eigenfunction associated to the minimum eigenvalue). One consequence is that many of the simplest smoothers (e.g., the scalar smoother) do not work for multigrid solutions to (1.2). In fact, other investigators have noticed this failure, either through analysis or computation, and have concluded that multigrid is not suitable for the problem (1.2). See, for example, [12]. However our analysis, which takes account of the special properties of the problem, shows that this conclusion is unjustified: with an appropriate smoother, multigrid is as efficient for the operator $\boldsymbol{\Lambda}$ as it is for a standard elliptic operator.

As indicated above, the operator $\boldsymbol{\Lambda}$ behaves very differently when applied to gradient fields than it does when applied to curl fields. This observation suggests that the Helmholtz decomposition of an arbitrary vectorfield into a gradient and a curl will provide insight. Particularly important in our analysis will be a discrete version of the Helmholtz decomposition that applies to functions in $\boldsymbol{V}_h$. However it is important to note that we use the discrete Helmholtz decomposition only as a theoretical tool: it is not necessary to compute it when applying our algorithms. In this respect our approach differs significantly from that of Vassilevski and Wang [33]. They also study multilevel preconditioners for the equation (1.2). However, their methods require the use of projections into spaces of curl fields at all levels, and this leads to algorithms with more complex structures.

After some preliminaries in § 2, in § 3 we introduce the finite element spaces we shall consider, namely the Raviart–Thomas spaces, and establish some new approximation properties for them. These results, which are intimately related to the discrete Helmholtz decomposition, will be crucial to the later analysis. § 4 and § 5 are devoted to the construction of domain decomposition and multigrid preconditioners for the discrete approximations of the operator $\boldsymbol{\Lambda}$, respectively. In § 6, we consider the extension of these results to related problems, namely when

the space $\boldsymbol{H}(\mathrm{div})$ is replaced by the subspace

$$\mathring{\boldsymbol{H}}(\mathrm{div}) = \{\, \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}) \,:\, \boldsymbol{v}\cdot\boldsymbol{n} = 0 \text{ on } \partial\Omega \,\},$$

and also when the inner product in $\boldsymbol{H}(\mathrm{div})$ is replaced by a parameter dependent inner product given by

$$\Lambda(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{u},\boldsymbol{v}) + k^2(\mathrm{div}\,\boldsymbol{u}, \mathrm{div}\,\boldsymbol{v}),$$

where $k \in (0,1]$. The applications of these results to mixed and least squares systems is given in § 7 and the results of some numerical computations are presented in § 8. In an appendix, we sketch the additional arguments that are needed to extend some of the results to the case of non-convex Ω, and in a second appendix we give the proofs of some basic lemmas from the theory of multigrid and domain decomposition.

2. Preliminaries

We suppose that the domain Ω is a convex polygon. For $S \subset \mathbb{R}^2$, we shall use the usual Sobolev spaces $H^m(S)$ with norm $\|\cdot\|_{m,S}$. The notation $\mathring{H}^1(S)$ is used for the subspace of functions in $H^1(S)$ which vanish on the boundary of S. When the set S coincides with Ω, we shall usually suppress it from the notation, and when the index m is zero, we shall usually suppress it. We use boldface type for vectors in $\mathbb{R}^2$, vector-valued functions, spaces of such functions, and operators with range in such spaces. Thus, for example, $\boldsymbol{L}^2$ denotes the space of 2-vector-valued functions on Ω for which both components are square integrable.

We shall use the standard differential operators

$$\mathbf{grad} = \begin{pmatrix}\partial/\partial x\\ \partial/\partial y\end{pmatrix}, \qquad \mathbf{curl} = \begin{pmatrix}-\partial/\partial y\\ \partial/\partial x\end{pmatrix}, \qquad \mathrm{div} = \begin{pmatrix}\partial/\partial x & \partial/\partial y\end{pmatrix}.$$

We shall study an additive Schwarz preconditioner in § 4 and an additive Schwarz smoother in § 5. Here we briefly recall the definition of the additive Schwarz operator in a general setting and some properties of it which we will need. For this purpose let V be a Hilbert space which can be decomposed into a finite (but not necessarily direct) sum of closed subspaces: $V = \sum_j V_j$. Let $B : V \to V$ be a symmetric positive definite operator and let $P_j : V \to V_j$ denote the orthogonal projection with respect to the norm $v \mapsto (Bv, v)^{1/2}$. The additive Schwarz operator may be written as $\Theta = \sum_j P_j B^{-1}$. It is easy to see that Θ is L^2-symmetric and positive definite. Moreover, for all $v \in V$,

$$(\Theta^{-1}v, v) = \inf_{\substack{v_j \in V_j \\ \sum_j v_j = v}} \sum_j (Bv_j, v_j). \tag{2.1}$$

For the convenience of the reader, we include a proof of this result in Appendix B.

3. Finite element discretizations

In this section we introduce the Raviart–Thomas finite element spaces. Let $\mathcal{T}_h$ be a quasiuniform family of triangulations of Ω, where $h > 0$ is a parameter representative of the diameter of the elements. For each non-negative integer r the Raviart–Thomas space of index r is given by

$$\boldsymbol{V}_h = \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}) : \boldsymbol{v}|_T \in \boldsymbol{P}_r(T) + (x,y)P_r(T) \text{ for all } T \in \mathcal{T}_h\}.$$

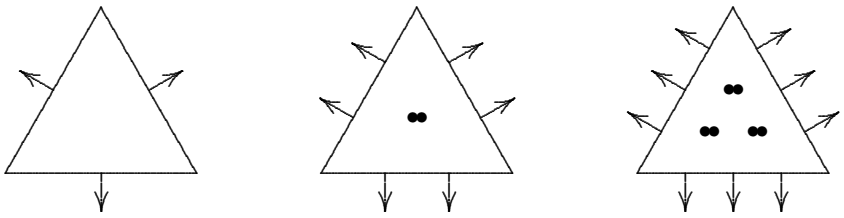

FIGURE 1. Degrees of freedom for the Raviart–Thomas spaces of indices 0, 1, and 2 approximating $\boldsymbol{H}(\mathrm{div})$. The arrows indicate the value of the normal component and the double dots the value of both components.

Here $P_r(T)$ denotes the set of polynomial functions of degree at most r on T. A vectorfield in $\boldsymbol{V}_h$ is uniquely specified by giving its value at a triangular array of $r(r+1)/2$ points in each triangle and the value of its normal component at $r+1$ points on each edge of the triangulation. Figure 1 shows the element diagram for $\boldsymbol{V}_h$ in the three lowest order cases.

We shall also introduce two other finite element spaces:

$$W_h = \{s \in H^1 : s|_T \in P_{r+1}(T)\}$$

is the usual space of continuous piecewise polynomials of degree $r+1$ and

$$S_h = \{q \in L^2 : q|_T \in P_r(T)\}$$

the space of arbitrary piecewise polynomials of degree r. It is easy to see that $\operatorname{div} \boldsymbol{V}_h \subset S_h$ and that $\mathbf{curl}\, W_h$ is precisely the subspace of divergence free vectorfields in $\boldsymbol{V}_h$ (cf. [10]):

$$\{\, \boldsymbol{v} \in \boldsymbol{V}_h : \ \operatorname{div} \boldsymbol{v} = 0 \,\} = \{\, \mathbf{curl}\, s : s \in W_h \,\}.$$

Defining the discrete gradient operator $\mathbf{grad}_h : S_h \to \boldsymbol{V}_h$ by the equation

$$(\mathbf{grad}_h\, q, \boldsymbol{v}) = -(q, \operatorname{div} \boldsymbol{v}), \ \text{for all } \boldsymbol{v} \in \boldsymbol{V}_h, \tag{3.1}$$

we immediately deduce the discrete Helmholtz decomposition (cf. [11])

$$\boldsymbol{V}_h = \mathbf{grad}_h\, S_h \oplus \mathbf{curl}\, W_h, \tag{3.2}$$

where the decomposition is orthogonal with respect to both the $\boldsymbol{L}^2$ and the $\boldsymbol{H}(\mathrm{div})$ inner products. Note that the two summand spaces $\mathbf{grad}_h\, S_h$ and $\mathbf{curl}\, W_h$ are invariant under the action of $\boldsymbol{\Lambda}_h$.

It is also well known (cf. [10] or [26]) that the pair of spaces $(\boldsymbol{V}_h, S_h)$ satisfies the inf-sup condition

$$\inf_{q \in S_h} \sup_{\boldsymbol{v} \in \boldsymbol{V}_h} \frac{(\operatorname{div} \boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div})} \|q\|} \geq \gamma > 0, \tag{3.3}$$

with the constant γ independent of h. The inf-sup condition follows from the existence of the interpolation operator $\boldsymbol{\Pi}_h : \boldsymbol{H}^1 \to \boldsymbol{V}_h$ having the commutativity property

$$\operatorname{div} \boldsymbol{\Pi}_h = Q_h \operatorname{div},$$

and the approximation property

$$\|\boldsymbol{u} - \boldsymbol{\Pi}_h \boldsymbol{u}\| \le ch\|\boldsymbol{u}\|_1, \text{ for all } \boldsymbol{u} \in \boldsymbol{H}^1.$$

Here $Q_h : L^2(\Omega) \to S_h$ is the L^2-projection onto S_h. In fact, the standard construction of $\boldsymbol{\Pi}_h$ determines it triangle by triangle from moments of $\boldsymbol{u} \cdot \boldsymbol{n}$ on the triangle edges and moments of $\boldsymbol{u}$ on the triangles. Therefore $\boldsymbol{\Pi}_h \boldsymbol{u}$ is defined for all $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div})$ for which $\boldsymbol{u}|_T \in \boldsymbol{H}^1(T)$ for all $T \in \mathcal{T}$, and for such $\boldsymbol{u}$,

$$\|\boldsymbol{u} - \boldsymbol{\Pi}_h \boldsymbol{u}\|_{0,T} \le ch\|\boldsymbol{u}\|_{1,T}.$$

Observe that (3.3) implies that $\mathbf{grad}_h$ is injective on S_h.

The Raviart–Thomas mixed method to approximate the solution of the Dirichlet problem

$$\Delta p = f \text{ on } \Omega, \quad p = 0 \text{ on } \partial\Omega, \tag{3.4}$$

determines $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times S_h$ by the equations

$$(\boldsymbol{u}_h, \boldsymbol{v}) + (\operatorname{div} \boldsymbol{v}, p_h) = 0 \text{ for all } \boldsymbol{v} \in \boldsymbol{V}_h,$$
$$(\operatorname{div} \boldsymbol{u}_h, q) = (f, q) \text{ for all } q \in S_h.$$

An equivalent formulation is

$$\boldsymbol{u}_h = \mathbf{grad}_h\, p_h, \quad \operatorname{div} \boldsymbol{u}_h = Q_h f.$$

The inf-sup condition implies that this is a stable numerical method in the sense that

$$\|\boldsymbol{u}_h\|_{\boldsymbol{H}(\mathrm{div})} + \|p_h\| \le C\|f\|$$

for some constant C independent of h and f. Moreover the following estimates are known for this method (see, e.g., [20]):

$$\|\boldsymbol{u} - \boldsymbol{u}_h\| \le \|\boldsymbol{u} - \boldsymbol{\Pi}_h \boldsymbol{u}\| \le ch^k \|\boldsymbol{u}\|_k, \quad k = 1, 2, \dots, r+1, \tag{3.5}$$

$$\|p - p_h\| \le ch^k \|p\|_k, \quad k = 2, 3, \dots, r+1, \tag{3.6}$$

$$\|p - p_h\| \le c(h\|p\|_1 + h^2\|p\|_2). \tag{3.7}$$

To close this section we consider the following situation. Suppose that a second quasiuniform mesh $\mathcal{T}_H$ of Ω with mesh size $H > h$ is given and define corresponding spaces $\boldsymbol{V}_H$, S_H, and W_H. Let $Q_H : L^2 \to S_H$ and $\boldsymbol{P}_H : \boldsymbol{H}(\mathrm{div}) \to \boldsymbol{V}_H$ denote the L^2- and $\boldsymbol{H}(\mathrm{div})$-projections, respectively. The following results, which will be crucial to the analysis in §§ 4 and 5, concerns the approximation of functions in S_h and $\boldsymbol{V}_h$ by those in S_H and $\boldsymbol{V}_H$.

Lemma 3.1. *Let $p_h \in S_h$, $\boldsymbol{v}_h = \mathbf{grad}_h\, p_h \in \boldsymbol{V}_h$. Define $p_H \in S_H$ and $\boldsymbol{v}_H \in \boldsymbol{V}_H$ by*

$$\boldsymbol{v}_H = \mathbf{grad}_H\, p_H, \qquad \operatorname{div} \boldsymbol{v}_H = Q_H \operatorname{div} \boldsymbol{v}_h.$$

Then

$$\|p_h - Q_H p_h\| \le cH\|\mathbf{grad}_h\, p_h\|, \tag{3.8}$$

$$\|\boldsymbol{v}_h - \boldsymbol{v}_H\| \le cH\|\operatorname{div} \boldsymbol{v}_h\|, \tag{3.9}$$

$$\|\boldsymbol{v}_h - \boldsymbol{v}_H\|_{\boldsymbol{H}(\mathrm{div})} \le cH\|\Lambda_h \boldsymbol{v}_h\|, \tag{3.10}$$

where the constant c is independent of h and H.

Proof. Define $p \in \mathring{H}^1(\Omega)$ by $\Delta p = \operatorname{div} \mathbf{grad}_h p_h$ and $\boldsymbol{v} = \mathbf{grad}\, p$. Then

$$\|p\|_1 \le c\| \mathbf{grad}_h p_h\|, \quad \|p\|_2 \le c\| \operatorname{div} \mathbf{grad}_h p_h\| \le ch^{-1}\| \mathbf{grad}_h p_h\|.$$

Moreover $(p_h, \boldsymbol{v}_h)$ is the mixed method approximation to $(p, \boldsymbol{v})$ in the space $S_h \times \boldsymbol{V}_h$ and $(p_H, \boldsymbol{v}_H)$ is the mixed method approximation to $(p, \boldsymbol{v})$ is $S_H \times \boldsymbol{V}_H$.

To establish the first inequality, we apply the triangle inequality, (3.7), the standard approximation property of Q_H, and the a priori estimates for p to write

$$\begin{aligned}\|p_h - Q_H p_h\| &\le \|p_h - p\| + \|Q_H(p - p_h)\| + \|p - Q_H p\| \\ &\le c(h\|p\|_1 + h^2\|p\|_2 + H\|p\|_1) \le cH\| \mathbf{grad}_h p_h\|.\end{aligned}$$

Applying (3.5) and noting that $\|\boldsymbol{v}\|_1 \le \|p\|_2 \le c\| \operatorname{div} \boldsymbol{v}_h\|$ gives

$$\|\boldsymbol{v} - \boldsymbol{v}_h\| \le ch\| \operatorname{div} \boldsymbol{v}_h\|, \qquad \|\boldsymbol{v} - \boldsymbol{v}_H\| \le cH\| \operatorname{div} \boldsymbol{v}_h\|,$$

and another application of the triangle inequality gives the second estimate.

Substituting $\operatorname{div} \boldsymbol{v}_h$ for p_h in the first estimate then gives

$$\| \operatorname{div} \boldsymbol{v}_h - \operatorname{div} \boldsymbol{v}_H\| \le cH\| \mathbf{grad}_h \operatorname{div} \boldsymbol{v}_h\|.$$

Combining this with the previous estimate gives

$$\|\boldsymbol{v}_h - \boldsymbol{v}_H\|_{\boldsymbol{H}(\mathrm{div})} \le cH(\| \operatorname{div} \boldsymbol{v}_h\| + \| \mathbf{grad}_h \operatorname{div} \boldsymbol{v}_h\|).$$

Since

$$\|\boldsymbol{\Lambda}_h \boldsymbol{v}_h\|^2 = \|\boldsymbol{v}_h\|^2 + 2\| \operatorname{div} \boldsymbol{v}_h\|^2 + \| \mathbf{grad}_h \operatorname{div} \boldsymbol{v}_h\|^2,$$

this establishes the third estimate. □

The key point in the next lemma is the gain of a power of H without bringing in the full $\boldsymbol{H}^1$ norm of $\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}$.

Lemma 3.1. *Suppose $\boldsymbol{u} \in \boldsymbol{V}_h$ and that $\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u} \in \boldsymbol{V}_h$ has the discrete Helmholtz decomposition*

$$\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u} = \mathbf{grad}_h p + \mathbf{curl}\, s,$$

for some $p \in S_h$ and $s \in W_h$. Then

$$\| \mathbf{grad}_h p\| + \|s\| \le cH\|\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}\|_{\boldsymbol{H}(\mathrm{div})},$$

where c is independent of H.

Proof. Clearly

$$\| \mathbf{grad}_h p\|^2 = \Lambda(\mathbf{grad}_h p, \boldsymbol{v}_h), \tag{3.11}$$

where $\boldsymbol{v}_h = \boldsymbol{\Lambda}_h^{-1} \mathbf{grad}_h p$. Since $\boldsymbol{\Lambda}_h$ maps $\mathbf{grad}_h S_h$ onto itself, we have $\boldsymbol{v}_h \in \mathbf{grad}_h S_h$, so

$$\Lambda(\mathbf{grad}_h p, \boldsymbol{v}_h) = \Lambda(\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}, \boldsymbol{v}_h). \tag{3.12}$$

Defining $\boldsymbol{v}_H \in \boldsymbol{V}_H$ as in the previous lemma then gives

$$\|\boldsymbol{v}_h - \boldsymbol{v}_H\|_{\boldsymbol{H}(\mathrm{div})} \le cH\|\Lambda_h \boldsymbol{v}_h\| = cH\| \mathbf{grad}_h p\|,$$

so

$$\begin{aligned}\Lambda(\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}, \boldsymbol{v}_h) &= \Lambda(\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}, \boldsymbol{v}_h - \boldsymbol{v}_H) \le \|\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}\|_{\boldsymbol{H}(\mathrm{div})}\|\boldsymbol{v}_h - \boldsymbol{v}_H\|_{\boldsymbol{H}(\mathrm{div})} \\ &\le cH\|\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}\|_{\boldsymbol{H}(\mathrm{div})}\| \mathbf{grad}_h p\|.\end{aligned} \tag{3.13}$$

The desired bound for $\|\mathbf{grad}_h\, p\|$ follows directly from (3.11), (3.12), and (3.13). Observe also that

$$(\mathbf{curl}\, s, \mathbf{curl}\, \psi) = \Lambda(\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}, \mathbf{curl}\, \psi) = 0 \text{ for all } \psi \in W_H.$$

Therefore, by a standard duality argument,

$$\|s\| \le cH \|\mathbf{curl}\, s\| \le cH \|\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}\|_{\boldsymbol{H}(\mathrm{div})}.$$

□

4. Domain decomposition methods

In this section we shall construct domain decomposition preconditioners for the operator $\boldsymbol{\Lambda}_h$ given by (1.1). We first define the additive Schwarz operator and show that it is an effective preconditioner by bounding its spectrum above and below. At the end of the section we define a multiplicative Schwarz operator and derive bounds on its spectrum from those for the additive operator. To the extent possible, we follow the standard analysis for second order elliptic operators (cf., for example, Dryja and Widlund [18], [19] or Xu [34]), but some modifications are required in order to handle the degeneracy of the operator.

Let $\mathcal{T}_H = \{\Omega_j\}_{j=1}^J$ be a coarse quasiuniform triangulation of Ω with characteristic grid size H, and let $\mathcal{T}_h$ be a quasiuniform refinement of $\mathcal{T}_H$ with characteristic grid size $h < H$. Furthermore, let $\{\Omega_j'\}_{j=1}^J$ be a covering of Ω such that for each j, $\bar{\Omega}_j'$ is a union of triangles in $\mathcal{T}_h$ and $\Omega_j \subset \Omega_j'$. It is also useful to set $\Omega_0 = \Omega_0' = \Omega$ and $\Gamma_j' = \partial\Omega_j' \setminus \partial\Omega$. We shall make the standard assumptions of bounded but sufficient overlap:

(A1) There is a constant β_1 such that each point of Ω is contained in at most β_1 of the sets Ω_j'.

(A2) There is a constant $\beta_2 > 0$ such that $\mathrm{dist}(\Gamma_j', \Omega_j) \ge \beta_2 H$.

Let $\boldsymbol{V}_h$, W_h, and S_h be the finite element spaces introduced in § 3, so that $\boldsymbol{V}_h$ is the Raviart–Thomas space of index r with respect to the triangulation $\mathcal{T}_h$ and the discrete Helmholtz decomposition (3.2) is satisfied. The same spaces formed with respect to the coarse mesh $\mathcal{T}_H$ will be denoted $\boldsymbol{V}_0$, W_0, and S_0. (It is also possible to use instead the Raviart-Thomas spaces of index 0 on the coarse level mesh without affecting any of the results we are about to establish.)

For $j = 1, 2, \ldots, J$ set

$$\begin{aligned}
\boldsymbol{V}_j &= \left\{ \boldsymbol{v} \in \boldsymbol{V} : \boldsymbol{v} \equiv 0 \text{ on } \Omega \setminus \Omega_j' \right\}, \\
W_j &= \left\{ w \in W : w \equiv 0 \text{ on } \Omega \setminus \Omega_j' \right\}, \\
S_j &= \left\{ q \in S : w \equiv 0 \text{ on } \Omega \setminus \Omega_j' \right\}.
\end{aligned}$$

Consider the domain Ω_j' with the triangulation induced by $\mathcal{T}_h$, and the Raviart–Thomas space of index r approximating $\boldsymbol{H}(\mathrm{div}, \Omega_j')$. Imposing the essential boundary condition $\boldsymbol{v} \cdot \boldsymbol{n} = 0$ on Γ_j' determines a subspace in which the degrees of freedom corresponding to nodes on Γ_j' are set equal to zero. We may imbed this subspace in $\boldsymbol{V}_h \subset \boldsymbol{H}(\mathrm{div})$ via extension by zero from Ω_j' to Ω and this is the space $\boldsymbol{V}_j$. Similarly, identifying functions on Ω_j' with their extension by zero to Ω, W_j is the usual Lagrangian finite element space of degree $r+1$ approximating the $H^1(\Omega_j')$ functions which vanish on Γ_j' and S_j is the usual space of discontinuous piecewise polynomials of degree r approximating $L^2(\Omega_j')$. We may then define a discrete gradient

operator $\mathbf{grad}_j : S_j \to \boldsymbol{V}_j$ as in (3.1) and the discrete Helmholtz decomposition $\boldsymbol{V}_j = \mathbf{grad}_j\, S_j + \mathbf{curl}\, W_j$ holds.

We use the decomposition $\boldsymbol{V}_h = \sum_{j=0}^{J} \boldsymbol{V}_j$ to define an additive Schwarz preconditioner, $\boldsymbol{\Theta}_h : \boldsymbol{V}_h \to \boldsymbol{V}_h$, as described in § 2. That is, $\boldsymbol{\Theta}_h = \sum_{j=0}^{J} \boldsymbol{P}_j \boldsymbol{\Lambda}_h^{-1}$, where $\boldsymbol{P}_j : \boldsymbol{V}_h \to \boldsymbol{V}_j$ is the $\boldsymbol{H}(\mathrm{div})$-orthogonal projection, or, in other words, $\boldsymbol{\Theta}_h \boldsymbol{f} = \sum_{j=0}^{J} \boldsymbol{u}_j$ where $\boldsymbol{u}_j \in \boldsymbol{V}_j$ solves the subdomain ($j > 0$) or coarse mesh ($j = 0$) problem

$$\Lambda(\boldsymbol{u}_j, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in \boldsymbol{V}_j.$$

To establish the effectiveness of this preconditioner, we need to provide bounds from above and below on the spectrum of

$$\boldsymbol{P} := \boldsymbol{\Theta}_h \boldsymbol{\Lambda}_h = \sum_{j=0}^{J} \boldsymbol{P}_j.$$

Such bounds are a direct consequence of the following theorem.

Theorem 4.1 . *There is a positive constant c (depending on the domain Ω, the overlap constants β_1 and β_2, and the shape and quasiuniformity constants for the meshes $\mathcal{T}_h$ and $\mathcal{T}_H$, but otherwise independent of h and H), such that*

$$c^{-1}\Lambda(\boldsymbol{u}, \boldsymbol{u}) \le \Lambda(\boldsymbol{P}\boldsymbol{u}, \boldsymbol{u}) \le c\Lambda(\boldsymbol{u}, \boldsymbol{u}) \qquad \textit{for all } \boldsymbol{u} \in \boldsymbol{V}_h.$$

Proof. First we shall establish the second inequality with $c = \beta_1$. By the Cauchy-Schwarz inequality, it is enough to show that

$$\Lambda(\boldsymbol{P}\boldsymbol{u}, \boldsymbol{P}\boldsymbol{u}) \le \beta_1 \Lambda(\boldsymbol{P}\boldsymbol{u}, \boldsymbol{u}). \tag{4.1}$$

Let χ_j be the characteristic function of Ω_j' so $\| \sum_{j=0}^{J} \chi_j^2 \|_{L^\infty} = \| \sum_{j=0}^{J} \chi_j \|_{L^\infty} \le \beta_1$. Since $\boldsymbol{P}\boldsymbol{u} = \sum_{j=0}^{J} \boldsymbol{P}_j \boldsymbol{u} = \sum_{j=0}^{J} \chi_j \boldsymbol{P}_j \boldsymbol{u}$ and $\operatorname{div} \boldsymbol{P}\boldsymbol{u} = \sum_{j=0}^{J} \chi_j \operatorname{div} \boldsymbol{P}_j \boldsymbol{u}$,

$$\begin{aligned}
\Lambda(\boldsymbol{P}\boldsymbol{u}, \boldsymbol{P}\boldsymbol{u}) &= \sum_{i,j=0}^{J} \int_\Omega \chi_i \chi_j [(\boldsymbol{P}_i \boldsymbol{u})(\boldsymbol{P}_j \boldsymbol{u}) + \operatorname{div}(\boldsymbol{P}_i \boldsymbol{u}) \operatorname{div}(\boldsymbol{P}_j \boldsymbol{u})]\, dx \\
&= \sum_{i,j=0}^{J} \int_\Omega (\chi_i \boldsymbol{P}_j \boldsymbol{u})(\chi_j \boldsymbol{P}_i \boldsymbol{u}) + [\chi_i \operatorname{div}(\boldsymbol{P}_j \boldsymbol{u})][\chi_j \operatorname{div}(\boldsymbol{P}_i \boldsymbol{u})]\, dx \\
&\le \sum_{i,j=0}^{J} \left\{ \int_\Omega \chi_i^2 [|\boldsymbol{P}_j \boldsymbol{u}|^2 + |\operatorname{div}(\boldsymbol{P}_j \boldsymbol{u})|^2]\, dx \right\}^{1/2} \left\{ \int_\Omega \chi_j^2 [|\boldsymbol{P}_i \boldsymbol{u}|^2 + |\operatorname{div}(\boldsymbol{P}_i \boldsymbol{u})|^2]\, dx \right\}^{1/2} \\
&\le \sum_{i,j=0}^{J} \int_\Omega \chi_i^2 [|\boldsymbol{P}_j \boldsymbol{u}|^2 + |\operatorname{div}(\boldsymbol{P}_j \boldsymbol{u})|^2]\, dx \\
&\le \beta_1 \sum_{j=0}^{J} \Lambda(\boldsymbol{P}_j \boldsymbol{u}, \boldsymbol{P}_j \boldsymbol{u}) = \beta_1 \sum_{j=0}^{J} \Lambda(\boldsymbol{P}_j \boldsymbol{u}, \boldsymbol{u}) = \beta_1 \Lambda(\boldsymbol{P}\boldsymbol{u}, \boldsymbol{u}),
\end{aligned}$$

as desired.

The key to establishing the first inequality is showing that we can decompose any $\boldsymbol{u} \in \boldsymbol{V}_h$ as $\sum_{j=0}^J \boldsymbol{u}_j$ with $\boldsymbol{u}_j \in \boldsymbol{V}_j$ and so that

$$\sum_{j=0}^J \Lambda(\boldsymbol{u}_j, \boldsymbol{u}_j) \le c\Lambda(\boldsymbol{u}, \boldsymbol{u}). \tag{4.2}$$

Assuming this result momentarily, the completion of the argument is standard:

$$\begin{aligned}\Lambda(\boldsymbol{u}, \boldsymbol{u}) &= \sum_{j=0}^J \Lambda(\boldsymbol{u}_j, \boldsymbol{u}) = \sum_{j=0}^J \Lambda(\boldsymbol{u}_j, \boldsymbol{P}_j \boldsymbol{u}) \\ &\le \left[\sum_{j=0}^J \Lambda(\boldsymbol{u}_j, \boldsymbol{u}_j)\right]^{1/2} \left[\sum_{j=0}^J \Lambda(\boldsymbol{P}_j \boldsymbol{u}, \boldsymbol{P}_j \boldsymbol{u})\right]^{1/2} \le [c\Lambda(\boldsymbol{u}, \boldsymbol{u})]^{1/2} \Lambda(\boldsymbol{P}\boldsymbol{u}, \boldsymbol{u})^{1/2}.\end{aligned}$$

To define the decomposition of $\boldsymbol{u}$, we first use the Helmholtz decomposition to write $\boldsymbol{u} = \mathbf{grad}_h\, p + \mathbf{curl}\, w$, with $p \in S_h$ and $w \in W_h$ normalized so that $\int w = 0$. We shall decompose each of the summands independently. The decomposition of $\mathbf{curl}\, w$ follows from the usual procedure to decompose $w \in H^1$ which arises in the theory of domain decomposition for standard elliptic operators (cf. [34, Lemma 7.1]). Using this procedure we may write w as $\sum_{j=0}^J w_j$ with $w_j \in W_j$ and $\sum_{j=0}^J \|w_j\|_1^2 \le c\|w\|_1^2$. Then $\mathbf{curl}\, w_j \in \boldsymbol{V}_j$, $\mathbf{curl}\, w = \sum_{j=0}^J \mathbf{curl}\, w_j$, and

$$\begin{aligned}\sum_{j=0}^J \Lambda(\mathbf{curl}\, w_j, \mathbf{curl}\, w_j) &\le \sum_{j=0}^J \|w_j\|_1^2 \le c\|w\|_1^2 \le c(\mathbf{curl}\, w, \mathbf{curl}\, w) \\ &= c\Lambda(\mathbf{curl}\, w, \mathbf{curl}\, w) \le c\Lambda(\boldsymbol{u}, \boldsymbol{u}).\end{aligned}$$

It thus remains to decompose $\boldsymbol{v} := \mathbf{grad}_h\, p$ as $\sum_{j=0}^J \boldsymbol{v}_j$ with $\boldsymbol{v}_j \in \boldsymbol{V}_j$ such that

$$\sum_{j=0}^J \Lambda(\boldsymbol{v}_j, \boldsymbol{v}_j) \le c\Lambda(\boldsymbol{v}, \boldsymbol{v}). \tag{4.3}$$

First define $(\boldsymbol{v}_0, p_0) \in \boldsymbol{V}_0 \times S_0$ by $\boldsymbol{v}_0 = \mathbf{grad}_0\, p_0$, $\operatorname{div} \boldsymbol{v}_0 = Q_0 \operatorname{div} \boldsymbol{v}$, where Q_0 is the L^2-projection into S_0. Next let $\{\theta_j\}_{j=1}^J$ be a partition of unity subordinate to the covering Ω_j' of Ω so that

$$\sum_{j=1}^J \theta_j \equiv 1, \qquad 0 \le \theta_j \le 1, \qquad \operatorname{supp}(\theta_j) \subset \Omega_j'.$$

In view of the sufficient overlap condition (A2), we can choose $\{\theta_j\}$ such that

$$\|\mathbf{grad}\, \theta_j\|_{\boldsymbol{L}^\infty} \le cH^{-1}. \tag{4.4}$$

We then set $\boldsymbol{v}_j = \boldsymbol{\Pi}_h[\theta_j(\boldsymbol{v} - \boldsymbol{v}_0)]$. By construction $\boldsymbol{v} = \sum_{j=0}^J \boldsymbol{v}_j$. Defining $\boldsymbol{u}_j = \mathbf{curl}\, w_j + \boldsymbol{v}_j$, we have

$$\sum_{j=0}^J \Lambda(\boldsymbol{u}_j, \boldsymbol{u}_j) \le 2 \sum_{j=0}^J \left[\Lambda(\mathbf{curl}\, w_j, \mathbf{curl}\, w_j) + \Lambda(\boldsymbol{v}_j, \boldsymbol{v}_j)\right].$$

Since $\Lambda(\boldsymbol{v},\boldsymbol{v}) \le \Lambda(\boldsymbol{u},\boldsymbol{u})$, the proof of the theorem will be complete if we can show (4.3). Now for $j > 0$,

$$\begin{aligned}\|\boldsymbol{v}_j\| &\le \|\boldsymbol{\Pi}_h[\theta_j(\boldsymbol{v}-\boldsymbol{v}_0)] - \theta_j(\boldsymbol{v}-\boldsymbol{v}_0)\| + \|\theta_j(\boldsymbol{v}-\boldsymbol{v}_0)\| \\ &\le ch\|\operatorname{\mathbf{grad}}_{\mathcal{T}_h}[\theta_j(\boldsymbol{v}-\boldsymbol{v}_0)]\| + \|\theta_j(\boldsymbol{v}-\boldsymbol{v}_0)\| \\ &\le ch[\|\operatorname{\mathbf{grad}}\theta_j\|_{\boldsymbol{L}^\infty}\|\boldsymbol{v}-\boldsymbol{v}_0\|_{\boldsymbol{L}^2(\Omega_j')} + \|\theta_j\|_{\boldsymbol{L}^\infty}\|\operatorname{\mathbf{grad}}_{\mathcal{T}_h}(\boldsymbol{v}-\boldsymbol{v}_0)\|_{\boldsymbol{L}^2(\Omega_j')}] \\ &\quad + \|\theta_j\|_{\boldsymbol{L}^\infty}\|\boldsymbol{v}-\boldsymbol{v}_0\|_{\boldsymbol{L}^2(\Omega_j')} \le c\|\boldsymbol{v}-\boldsymbol{v}_0\|_{\boldsymbol{L}^2(\Omega_j')}.\end{aligned}$$

Here $\operatorname{\mathbf{grad}}_{\mathcal{T}_h}$ denotes the gradient applied elementwise with respect to the elements in $\mathcal{T}_h$, and we have used an inverse inequality, the boundedness of θ_j, and (4.4) in the last step.

Now

$$\begin{aligned}\|\operatorname{div}\boldsymbol{\Pi}_h[\theta_j(\boldsymbol{v}-\boldsymbol{v}_0)]\| &\le \|\operatorname{div}[\theta_j(\boldsymbol{v}-\boldsymbol{v}_0)]\| = \|\operatorname{div}[\theta_j(\boldsymbol{v}-\boldsymbol{v}_0)]\|_{0,\Omega_j'} \\ &\le \|\operatorname{\mathbf{grad}}\theta_j\|_{\boldsymbol{L}^\infty}\|\boldsymbol{v}-\boldsymbol{v}_0\|_{0,\Omega_j'} + \|\theta_j\|_{\boldsymbol{L}^\infty}\|\operatorname{div}(\boldsymbol{v}-\boldsymbol{v}_0)\|_{0,\Omega_j'} \\ &\le c(H^{-1}\|\boldsymbol{v}-\boldsymbol{v}_0\|_{0,\Omega_j'} + \|\operatorname{div}(\boldsymbol{v}-\boldsymbol{v}_0)\|_{0,\Omega_j'}).\end{aligned}$$

Squaring and adding over j, and using (A1) gives

$$\sum_{j=1}^{J}\Lambda(\boldsymbol{v}_j,\boldsymbol{v}_j) \le c[(1+H^{-2})\|\boldsymbol{v}-\boldsymbol{v}_0\|^2 + \|\operatorname{div}(\boldsymbol{v}-\boldsymbol{v}_0)\|^2]. \tag{4.5}$$

Since $\operatorname{div}\boldsymbol{v}_0 = Q_0\operatorname{div}\boldsymbol{v}$, it follows easily using Lemma 3.1 that the right-hand side may be bounded by $c\Lambda(\boldsymbol{v},\boldsymbol{v})$. This completes the proof. □

We remark that in using Lemma 3.1, we require that the domain Ω be convex. In fact, this restriction may be eliminated by using a more complicated argument, which we present in Appendix A.

We shall now discuss the corresponding symmetric multiplicative Schwarz preconditioner $\widetilde{\boldsymbol{\Theta}}_h : \boldsymbol{V}_h \to \boldsymbol{V}_h$. The results for this preconditioner follow by a standard argument from those derived above for the additive preconditioner (cf. for example Bramble, Pasciak, Wang, and Xu [7] or Xu [34]). We recall the definition of the multiplicative Schwarz operator $\widetilde{\boldsymbol{\Theta}}_h$. For a given $\boldsymbol{f} \in \boldsymbol{V}_h$ we let $\widetilde{\boldsymbol{\Theta}}_h\boldsymbol{f} = \boldsymbol{v}^J \in \boldsymbol{V}_h$, where the $\boldsymbol{v}^j$ are defined by the iteration

$$\begin{aligned}&\boldsymbol{v}^{-J-1} = 0, \\ &\boldsymbol{v}^j = \boldsymbol{v}^{j-1} - \boldsymbol{P}_{|j|}(\boldsymbol{v}^{j-1} - \boldsymbol{\Lambda}_h^{-1}\boldsymbol{f}), \quad j = -J, -J+1, \dots, J.\end{aligned} \tag{4.6}$$

If we let $\boldsymbol{v} = \boldsymbol{\Lambda}_h^{-1}\boldsymbol{f}$, then it follows from the iteration above that

$$\boldsymbol{v} - \boldsymbol{v}^j = (\boldsymbol{I} - \boldsymbol{P}_{|j|})(\boldsymbol{v} - \boldsymbol{v}^{j-1}).$$

Hence, if we let $\boldsymbol{E} : \boldsymbol{V}_h \to \boldsymbol{V}_h$ denote the operator

$$\boldsymbol{E} = (\boldsymbol{I} - \boldsymbol{P}_0)(\boldsymbol{I} - \boldsymbol{P}_1)\dots(\boldsymbol{I} - \boldsymbol{P}_J),$$

then

$$\boldsymbol{I} - \widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h = \boldsymbol{E}^*\boldsymbol{E},$$

where $\boldsymbol{E}^*$ is the adjoint of $\boldsymbol{E}$ with respect to the inner product $\Lambda(\cdot,\cdot)$. This immediately implies that

$$\Lambda([\boldsymbol{I} - \widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h]\boldsymbol{u},\boldsymbol{u}) = \Lambda(\boldsymbol{E}\boldsymbol{u},\boldsymbol{E}\boldsymbol{u}) \ge 0 \qquad \text{for all } \boldsymbol{u} \in \boldsymbol{V}_h, \tag{4.7}$$

and hence the spectrum of $\widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h$ is bounded above by one.

A lower bound on the spectrum of $\widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h$ can be derived from an upper bound of $\Lambda(\boldsymbol{E}\boldsymbol{u},\boldsymbol{E}\boldsymbol{u})$. Define operators

$$\boldsymbol{E}_j = (\boldsymbol{I}-\boldsymbol{P}_j)(\boldsymbol{I}-\boldsymbol{P}_{j+1})\dots(\boldsymbol{I}-\boldsymbol{P}_J) \qquad \text{for } j=0,1,\dots J.$$

Hence $\boldsymbol{E}_0=\boldsymbol{E}$. Furthermore let $\boldsymbol{E}_{J+1}=\boldsymbol{I}$. Observe that

$$\begin{aligned}\Lambda(\boldsymbol{E}_j\boldsymbol{u},\boldsymbol{E}_j\boldsymbol{u}) &= \Lambda([\boldsymbol{I}-\boldsymbol{P}_j]\boldsymbol{E}_{j+1}\boldsymbol{u},\boldsymbol{E}_{j+1}\boldsymbol{u})\\ &= \Lambda(\boldsymbol{E}_{j+1}\boldsymbol{u},\boldsymbol{E}_{j+1}\boldsymbol{u})-\Lambda(\boldsymbol{P}_j\boldsymbol{E}_{j+1}\boldsymbol{u},\boldsymbol{E}_{j+1}\boldsymbol{u}).\end{aligned}$$

Therefore we obtain

$$\Lambda(\boldsymbol{E}\boldsymbol{u},\boldsymbol{E}\boldsymbol{u}) = \Lambda(\boldsymbol{u},\boldsymbol{u}) - \sum_{j=0}^{J}\Lambda(\boldsymbol{P}_j\boldsymbol{E}_{j+1}\boldsymbol{u},\boldsymbol{E}_{j+1}\boldsymbol{u}).$$

Combining this with (4.7) we have

$$\Lambda(\widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h\boldsymbol{u},\boldsymbol{u}) = \sum_{j=0}^{J}\Lambda(\boldsymbol{P}_j\boldsymbol{E}_{j+1}\boldsymbol{u},\boldsymbol{E}_{j+1}\boldsymbol{u}). \tag{4.8}$$

The desired lower bound for the spectrum of $\widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h$ will essentially follow from the identity (4.8) and the corresponding bound for the additive operator derived above. In order to see this, observe first that from the relation $\boldsymbol{E}_j=(\boldsymbol{I}-\boldsymbol{P}_j)\boldsymbol{E}_{j+1}$ we obtain

$$\boldsymbol{I} = \boldsymbol{E}_j + \sum_{i=j}^{J}\boldsymbol{P}_i\boldsymbol{E}_{i+1}.$$

Therefore, since $\boldsymbol{P}_j\boldsymbol{E}_j=0$ we have

$$\Lambda(\boldsymbol{P}\boldsymbol{u},\boldsymbol{u}) = \sum_{j=0}^{J}\Lambda(\boldsymbol{P}_j\boldsymbol{u},\boldsymbol{u}) = \sum_{j=0}^{J}\sum_{i=j}^{J}\Lambda(\boldsymbol{P}_j\boldsymbol{u},\boldsymbol{P}_i\boldsymbol{E}_{i+1}\boldsymbol{u}). \tag{4.9}$$

By arguing as in the derivation of (4.1), it now follows from the Cauchy–Schwarz inequality and the overlap condition (A1) that

$$\sum_{j=0}^{J}\sum_{i=j}^{J}\Lambda(\boldsymbol{P}_j\boldsymbol{u},\boldsymbol{P}_i\boldsymbol{E}_{i+1}\boldsymbol{u}) \le \beta_1\Lambda(\boldsymbol{P}\boldsymbol{u},\boldsymbol{u})^{1/2}\Big[\sum_{j=0}^{J}\Lambda(\boldsymbol{P}_j\boldsymbol{E}_{j+1}\boldsymbol{u},\boldsymbol{E}_{j+1}\boldsymbol{u})\Big]^{1/2}.$$

Together with (4.9) this implies that

$$\Lambda(\boldsymbol{P}\boldsymbol{u},\boldsymbol{u}) \le \beta_1^2\sum_{j=0}^{J}\Lambda(\boldsymbol{P}_j\boldsymbol{E}_{j+1}\boldsymbol{u},\boldsymbol{E}_{j+1}\boldsymbol{u}). \tag{4.10}$$

Hence, by combining (4.8), (4.10), and the left inequality in Theorem 4.1, we obtain

$$\Lambda(\widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h\boldsymbol{u},\boldsymbol{u}) \ge \beta_1^{-2}\Lambda(\boldsymbol{P}\boldsymbol{u},\boldsymbol{u}) \ge (c\beta_1^2)^{-1}\Lambda(\boldsymbol{u},\boldsymbol{u}) \qquad \text{for all } \boldsymbol{u}\in\boldsymbol{V}_h. \tag{4.11}$$

We summarize this discussion of the multiplicative Schwarz operator in the following theorem.

Theorem 4.2 . *The spectrum of the operator $\widetilde{\boldsymbol{\Theta}}_h\boldsymbol{\Lambda}_h$ is contained in an interval $[1-\delta,1]$, where the positive constant δ is independent of h and H, but depends on the overlap constants β_1 and β_2.*

As a corollary of the discussion which led to Theorem 4.2, we also obtain the following relation between the multiplicative operator $\widetilde{\boldsymbol{\Theta}}_h$ and the corresponding additive operator $\boldsymbol{\Theta}_h$ which will be used in the next section.

Corollary 4.3. *Let $\beta_1 > 0$ be given by the overlap condition (A1). Then*

$$(\widetilde{\boldsymbol{\Theta}}_h \boldsymbol{v}, \boldsymbol{v}) \geq \beta_1^{-2} (\boldsymbol{\Theta}_h \boldsymbol{v}, \boldsymbol{v}) \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{V}_h.$$

Proof. From (4.11) we have

$$(\widetilde{\boldsymbol{\Theta}}_h \boldsymbol{\Lambda}_h \boldsymbol{u}, \boldsymbol{\Lambda}_h \boldsymbol{u}) \geq \beta_1^{-2} (\boldsymbol{\Theta}_h \boldsymbol{\Lambda}_h \boldsymbol{u}, \boldsymbol{\Lambda}_h \boldsymbol{u}) \qquad \text{for all } \boldsymbol{u} \in \boldsymbol{V}_h.$$

The result follows by setting $\boldsymbol{v} = \boldsymbol{\Lambda}_h^{-1} \boldsymbol{u}$. □

5. Multigrid methods

In this section we define a V–cycle multigrid preconditioner $\boldsymbol{\Theta}_h$ for the operator $\boldsymbol{\Lambda}_h$ using an additive Schwarz smoother formed by summing solutions to local problems in a neighborhood of each mesh vertex, and we show that the operator $\boldsymbol{I} - \boldsymbol{\Theta}_h \boldsymbol{\Lambda}_h$ is a contraction uniformly with respect to h, and, *a fortiori*, that $\boldsymbol{\Theta}_h$ is spectrally equivalent to $\boldsymbol{\Lambda}_h^{-1}$. At the end of this section, we show that an analogous result holds for the multiplicative Schwarz smoother.

We begin by recalling the multigrid V-cycle construction in an abstract setting. For this discussion the notation is independent of the rest of the paper. Let $V_1 \subset V_2 \subset \ldots \subset V_J$ be a nested sequence of finite dimensional subspaces of a Hilbert space H, and let $\Lambda : V_J \times V_J \to \mathbb{R}$ be a positive-definite symmetric bilinear form. For $j = 1, 2, \ldots, J$ define $\Lambda_j : V_j \to V_j$ by

$$(\Lambda_j v, w) = \Lambda(v, w) \text{ for all } v, w \in V_j,$$

where the pairing on the left-hand side is the inner product in H. Also let $Q_j : V_J \to V_j$ denote the H-orthogonal projection and $P_j : V_J \to V_j$ the orthogonal projection with respect to the bilinear form Λ. Finally, suppose that we are given for each $j > 1$ a linear operator $R_j : V_j \to V_j$. As will be clarified below, R_j, the smoother, is intended to behave in some ways like an approximation to Λ_j^{-1}.

For a fixed positive integer m, the standard V–cycle multigrid algorithm with m smoothings recursively defines operators $\Theta_j : V_j \to V_j$ beginning with $\Theta_1 = \Lambda_1^{-1}$. For $j > 1$ and $f \in V_j$ we define $\Theta_j f = x_{2m+1}$ where

$$\begin{aligned} x_0 &= 0 \in V_j, \\ x_i &= x_{i-1} + R_j(f - \Lambda_j x_{i-1}), \quad i = 1, 2, \ldots, m, \\ x_{m+1} &= x_m + \Theta_{j-1} Q_{j-1}(f - \Lambda_j x_m), \\ x_i &= x_{i-1} + R_j(f - \Lambda_j x_{i-1}), \quad i = m+2, m+3, \ldots, 2m+1. \end{aligned}$$

Note that if R_j is H-symmetric, as we shall assume, then so is Θ_j.

The following result is useful for establishing the convergence of the V–cycle algorithm. It was proved in a special case by Braess and Hackbusch [2] and can easily be adapted from the proof of Theorem 3.6 in [3]. For the convenience of the reader, we provide a proof in Appendix B.

Theorem 5.1. *Suppose that for each $j = 1, 2, \ldots, J$ the smoother R_j is H-symmetric and positive semidefinite and satisfies the conditions*

$$\Lambda([I - R_j \Lambda_j] v, v) \geq 0 \text{ for all } v \in V_j \tag{5.1}$$

and

(5.2)
$$(R_j^{-1}[I - P_{j-1}]v, [I - P_{j-1}]v) \le \alpha\Lambda([I - P_{j-1}]v, [I - P_{j-1}]v) \text{ for all } v \in V_j,$$

where α is some constant. Then

$$0 \le \Lambda([I - \Theta_j\Lambda_j]v, v) \le \delta\Lambda(v, v) \text{ for all } v \in V_j,$$

where $\delta = \alpha/(\alpha + 2m)$.

Corollary 5.2 . *Under the hypotheses of the theorem, the error operator $I - \Theta_J\Lambda_J$ is a positive definite contraction on V_J whose operator norm relative to the Λ inner product is bounded by δ. Moreover the eigenvalues of $\Theta_J\Lambda_J$ belong to the interval $[1 - \delta, 1]$.*

Thus we have, in particular, that Θ_J is spectrally equivalent to Λ_J^{-1}.

We wish to apply this abstract theorem to the case where $H = \boldsymbol{L}^2$, $V_J = \boldsymbol{V}_h$ is the Raviart–Thomas space of index $r \ge 0$ relative to the triangulation $\mathcal{T}_h$, and Λ is the $\boldsymbol{H}(\mathrm{div})$ inner product. In order to define the nested sequence of subspaces V_j, we assume that the triangulation $\mathcal{T}_h$ is constructed by a successive refinement process. More precisely, we assume that we have a nested sequence of quasiuniform triangulations $\mathcal{T}_j$, $1 \le j \le J$, with characteristic mesh size h_j proportional to γ^{2j} for some positive constant $\gamma < 1$, and that $\mathcal{T}_h = \mathcal{T}_J$. It is easy to check that

$$\boldsymbol{V}_1 \subset \boldsymbol{V}_2 \subset \cdots \subset \boldsymbol{V}_J = \boldsymbol{V}_h$$

where $\boldsymbol{V}_j$ is the Raviart-Thomas space of index r relative to the triangulation $\mathcal{T}_j$. (Note, that $\boldsymbol{V}_j$ in this section has a different meaning than it had in the preceding one.) At each level j we have the discrete operator $\boldsymbol{\Lambda}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$ defined as in (1.1) and the $\boldsymbol{L}^2$- and $\boldsymbol{H}(\mathrm{div})$-orthogonal projections onto $\boldsymbol{V}_j$, which we denote by $\boldsymbol{Q}_j$ and $\boldsymbol{P}_j$, respectively. Also, using the triangulation $\mathcal{T}_j$ we may define the spaces W_j and S_j and the discrete gradient operator $\mathbf{grad}_j : S_j \to \boldsymbol{V}_j$ so that the discrete Helmholtz decomposition

$$\boldsymbol{V}_j = \mathbf{grad}_j\, S_j \oplus \mathbf{curl}\, W_j,$$

holds. Thus there exist maps $F_j : \boldsymbol{V}_j \to W_j$ and $G_j : \boldsymbol{V}_j \to S_j$, so that

$$\boldsymbol{u} = \mathbf{grad}_j(G_j\boldsymbol{u}) + \mathbf{curl}(F_j\boldsymbol{u}) \text{ for all } \boldsymbol{u} \in \boldsymbol{V}_j.$$

To complete the description of the multigrid algorithm, we must define the smoothers. We shall discuss both additive smoothers $\boldsymbol{R}_j$ and multiplicative smoothers $\widetilde{\boldsymbol{R}}_j$. The additive smoother $\boldsymbol{R}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$ will be defined as a multiple of the additive Schwarz operator formed with respect to a decomposition of $\boldsymbol{V}_j$ which we now describe. Let $\mathcal{N}_j$ be the set of vertices in the triangulation $\mathcal{T}_j$, and for each $\nu \in \mathcal{N}_j$ let $\mathcal{T}_{j,\nu}$ be the set of triangles in $\mathcal{T}_j$ meeting at the vertex ν. These form a triangulation of a small subdomain which we denote $\Omega_{j,\nu}$. The family of subdomains $\{\Omega_{j,\nu}\}_{\nu\in\mathcal{N}_j}$ forms an overlapping covering of Ω as did the family $\{\Omega'_j\}_{j=1}^J$ of the preceding section, and we define a decomposition of $\boldsymbol{V}_j$, W_j, and S_j as we did there: $\boldsymbol{V}_{j,\nu}$, $W_{j,\nu}$, and $S_{j,\nu}$ are the subsets of functions in $\boldsymbol{V}_j$, W_j, and S_j, respectively, which are supported in $\bar{\Omega}_{j,\nu}$. Note that the finite overlap condition certainly holds: no point belongs to more than three of the $\Omega_{j,\nu}$. Then we define

$$\boldsymbol{R}_j = \eta \sum_{\nu\in\mathcal{N}_j} \boldsymbol{P}_{j,\nu}\Lambda_j^{-1} \tag{5.3}$$

as the additive Schwarz operator multiplied by a scaling factor of $\eta > 0$. The implementation of $\boldsymbol{R}_j$ is easy: on each of the small domains (each consisting of only a small number of elements, whose number does not grow with decreasing h or increasing j) one must solve the restricted discrete variational problem with Dirichlet boundary conditions on the subdomain boundary (except where the subdomain boundary coincides with the domain boundary), and the scaled sum of these solutions gives the value of $\boldsymbol{R}_j$. As discussed at the end of § 2, $\boldsymbol{R}_j$ is L^2-symmetric and positive definite and

$$\eta(\boldsymbol{R}_j^{-1}\boldsymbol{v},\boldsymbol{v}) = \inf_{\substack{\boldsymbol{v}_\nu\in\boldsymbol{V}_{j,\nu}\\ \sum_\nu \boldsymbol{v}_\nu=\boldsymbol{v}}} \sum_{\nu\in\mathcal{N}_j} \Lambda(\boldsymbol{v}_\nu,\boldsymbol{v}_\nu). \tag{5.4}$$

In the remainder of this section we show that if $0 < \eta \le 1/3$, then the scaled additive Schwarz smoother (5.3) satisfies the conditions of Theorem 5.1 and so the conclusions of that theorem and of Corollary 5.2 are satisfied.

For $\boldsymbol{v} \in \boldsymbol{V}_j$,

$$\Lambda([\boldsymbol{I} - \boldsymbol{R}_j\boldsymbol{\Lambda}_j]\boldsymbol{v},\boldsymbol{v}) = \Lambda(\boldsymbol{v},\boldsymbol{v}) - \eta \sum_{\nu\in\mathcal{N}_j} \Lambda(\boldsymbol{P}_{j,\nu}\boldsymbol{v},\boldsymbol{v}).$$

But

$$\Lambda(\boldsymbol{P}_{j,\nu}\boldsymbol{v},\boldsymbol{v}) = \|\boldsymbol{P}_{j,\nu}\boldsymbol{v}\|^2_{\boldsymbol{H}(\mathrm{div},\Omega_{j,\nu})} \le \|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div},\Omega_{j,\nu})}\|\boldsymbol{P}_{j,\nu}\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div},\Omega_{j,\nu})},$$

so

$$\sum_{\nu\in\mathcal{N}_j} \Lambda(\boldsymbol{P}_{j,\nu}\boldsymbol{v},\boldsymbol{v}) \le \sum_{\nu\in\mathcal{N}_j} \|\boldsymbol{v}\|^2_{\boldsymbol{H}(\mathrm{div},\Omega_{j,\nu})} \le 3\Lambda(\boldsymbol{v},\boldsymbol{v}).$$

Thus the hypothesis (5.1) holds.

Thus, it only remains to establish (5.2), which, in view of (5.4) reduces to showing that for $\boldsymbol{v} = (\boldsymbol{I} - \boldsymbol{P}_{j-1})\boldsymbol{u}$, $\boldsymbol{u} \in \boldsymbol{V}_j$, we can decompose $\boldsymbol{v}$ as $\sum_\nu \boldsymbol{v}_\nu$ with $\boldsymbol{v}_\nu \in \boldsymbol{V}_{j,\nu}$ such that

$$\sum_{\nu\in\mathcal{N}_j} \Lambda(\boldsymbol{v}_\nu,\boldsymbol{v}_\nu) \le c\Lambda(\boldsymbol{v},\boldsymbol{v}). \tag{5.5}$$

We use the discrete Helmholtz decomposition to write

$$\boldsymbol{v} = \bar{\boldsymbol{v}} + \tilde{\boldsymbol{v}} := \mathbf{grad}_h\, G_j\boldsymbol{v} + \mathbf{curl}\, F_j\boldsymbol{v},$$

and decompose the two pieces separately.

First we consider $\bar{\boldsymbol{v}} := \mathbf{grad}_h\, G_j\boldsymbol{v}$. Letting $\{\theta_\nu\}_{\nu\in\mathcal{N}_j}$ denote a partition of unity subordinate to the covering $\{\Omega_{j,\nu}\}_{\nu\in\mathcal{N}_j}$, we set $\bar{\boldsymbol{v}}_\nu = \boldsymbol{\Pi}_j(\theta_\nu\bar{\boldsymbol{v}})$. Then $\bar{\boldsymbol{v}} = \sum_\nu \bar{\boldsymbol{v}}_\nu$, and, arguing as at the end of the proof of Theorem 4.1 (with $H = h_j$ and $\boldsymbol{v}_0 = 0$), we get

$$\sum_{\nu\in\mathcal{N}_j} \Lambda(\bar{\boldsymbol{v}}_\nu,\bar{\boldsymbol{v}}_\nu) \le C(\|\bar{\boldsymbol{v}}\|^2_{\boldsymbol{H}(\mathrm{div})} + h_j^{-2}\|\bar{\boldsymbol{v}}\|^2). \tag{5.6}$$

Clearly $\|\bar{\boldsymbol{v}}\|_{\boldsymbol{H}(\mathrm{div})} \le \|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div})}$ and by Lemma 3.2, $\|\bar{\boldsymbol{v}}\| \le ch_j\|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div})}$, so

$$\sum_{\nu\in\mathcal{N}_j} \Lambda(\bar{\boldsymbol{v}}_\nu,\bar{\boldsymbol{v}}_\nu) \le C\|\boldsymbol{v}\|^2_{\boldsymbol{H}(\mathrm{div})}.$$

Next consider $\tilde{\boldsymbol{v}} = \mathbf{curl}\, F_j \boldsymbol{v}$. Using a standard decomposition argument (as in [34, Lemma 7.1]) we may write $F_j \boldsymbol{v} = \sum_\nu s_\nu$, where $s_\nu \in W_{j,\nu}$ and

$$\sum_{\nu \in \mathcal{N}_j} (\mathbf{curl}\, s_\nu, \mathbf{curl}\, s_\nu) \le C[(\mathbf{curl}\, F_j \boldsymbol{v}, \mathbf{curl}\, F_j \boldsymbol{v}) + h_j^{-2} \|F_j \boldsymbol{v}\|^2].$$

Recalling that $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{P}_{j-1}\boldsymbol{u} = \mathbf{grad}_h\, G_j \boldsymbol{v} + \mathbf{curl}\, F_j \boldsymbol{v}$, we have from Lemma 3.2 that

$$\|F_j \boldsymbol{v}\| \le C h_j \Lambda(\boldsymbol{v}, \boldsymbol{v})^{1/2}.$$

Setting $\tilde{\boldsymbol{v}}_\nu = \mathbf{curl}\, s_\nu$ so $\tilde{\boldsymbol{v}} = \sum_\nu \tilde{\boldsymbol{v}}_\nu$, and combining these results, we obtain

$$\sum_{\nu \in \mathcal{N}_j} \Lambda(\tilde{\boldsymbol{v}}_\nu, \tilde{\boldsymbol{v}}_\nu) = \sum_{\nu \in \mathcal{N}_j} (\mathbf{curl}\, s_\nu, \mathbf{curl}\, s_\nu) \le C \Lambda(\boldsymbol{v}, \boldsymbol{v}).$$

Finally, setting $\boldsymbol{v}_\nu = \bar{\boldsymbol{v}}_\nu + \tilde{\boldsymbol{v}}_\nu$, we get that $\boldsymbol{v} = \sum \boldsymbol{v}_\nu$ and that (5.5) holds. This completes the verification of the required properties of the smoother.

Next we consider the corresponding multiplicative preconditioner $\widetilde{\boldsymbol{R}}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$. This operator is defined by an algorithm of the form (4.6) with respect to the spaces $\boldsymbol{V}_{j,\nu}$. By construction (cf. (4.7)), these operators satisfy (5.1). Furthermore, by Corollary 4.3, the operators $\widetilde{\boldsymbol{R}}_j$ and $\boldsymbol{R}_j$ satisfy

$$(\boldsymbol{R}_j \boldsymbol{v}, \boldsymbol{v}) \le 9\eta (\widetilde{\boldsymbol{R}}_j \boldsymbol{v}, \boldsymbol{v}) \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{V}_h. \tag{5.7}$$

In order to verify (5.2) for the multiplicative operator, again let $\boldsymbol{v} = (\boldsymbol{I} - \boldsymbol{P}_{j-1})\boldsymbol{u}$, $\boldsymbol{u} \in \boldsymbol{V}_j$, and decompose $\boldsymbol{v}$ as $\sum_\nu \boldsymbol{v}_\nu$, with $\boldsymbol{v}_\nu \in \boldsymbol{V}_{j,\nu}$ satisfying (5.5). Then, by the Cauchy–Schwarz inequality and the definition of the additive smoother $\boldsymbol{R}_j$, we have

$$\begin{aligned}
(\widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}, \boldsymbol{v}) &= \sum_{\nu \in \mathcal{N}_j} \Lambda(\boldsymbol{P}_{j,\nu} \boldsymbol{\Lambda}_j^{-1} \widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}, \boldsymbol{v}_\nu) \\
&\le \Big[\sum_{\nu \in \mathcal{N}_j} \Lambda(\boldsymbol{P}_{j,\nu} \boldsymbol{\Lambda}_j^{-1} \widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}, \boldsymbol{\Lambda}_j^{-1} \widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}) \Big]^{1/2} \Big[\sum_{\nu \in \mathcal{N}_j} \Lambda(\boldsymbol{v}_\nu, \boldsymbol{v}_\nu) \Big]^{1/2} \\
&= \eta^{-1/2} (\boldsymbol{R}_j \widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}, \widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v})^{1/2} \Big[\sum_{\nu \in \mathcal{N}_j} \Lambda(\boldsymbol{v}_\nu, \boldsymbol{v}_\nu) \Big]^{1/2}.
\end{aligned}$$

Hence, from (5.5) and (5.7) we obtain

$$(\widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}, \boldsymbol{v}) \le c (\widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}, \boldsymbol{v})^{1/2} \Lambda(\boldsymbol{v}, \boldsymbol{v})^{1/2}$$

which implies (5.2).

6. Extensions

We now consider two extensions of the results obtained in the previous sections. First we remark that the entire analysis of §§ 3–5 adapts easily to the case where the space $\boldsymbol{H}(\mathrm{div})$ is replaced by the subspace

$$\mathring{\boldsymbol{H}}(\mathrm{div}) = \{\, \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}) \,:\, \boldsymbol{v} \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega \,\}.$$

In this case, the finite element spaces $\boldsymbol{V}_h$, W_h, and S_h are replaced by $\boldsymbol{V}_h \cap \mathring{\boldsymbol{H}}(\mathrm{div})$, $W_h \cap \mathring{H}^1$, and $S_h \cap \hat{L}^2$, respectively, and the Dirichlet problem (3.4) is replaced by

the Neumann problem

$$\Delta p = f \text{ on } \Omega, \quad \partial p/\partial n = 0 \text{ on } \partial\Omega, \quad \int_\Omega p\,dx = 0.$$

(Here $\hat{L}^2$ denotes the subspace of functions in L^2 with mean value zero.)

Second, we consider the case where the bilinear form Λ is redefined as

$$\Lambda(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}) + k^2(\operatorname{div} \boldsymbol{u}, \operatorname{div} \boldsymbol{v}), \tag{6.1}$$

where $k \in (0, 1]$ is a parameter. (An application of this case will be discussed in the next section.) The results of §§ 4 and 5 continue to hold in this case, moreover, with the spectral bounds now independent of k as well as h. In order to prove such results, some small modifications to the preceding analysis are required. We now discuss these modifications. First we define the norm

$$\|\boldsymbol{v}\|_\Lambda = [\Lambda(\boldsymbol{v}, \boldsymbol{v})]^{1/2}.$$

In the case when $k = 1$, this is the norm $\|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div})}$. The first change we need occurs in Lemma 3.1, where we need to replace the estimate

$$\|\boldsymbol{v}_h - \boldsymbol{v}_H\|_{\boldsymbol{H}(\mathrm{div})} \le cH\|\Lambda_h \boldsymbol{v}_h\|$$

by the estimate

$$\|\boldsymbol{v}_h - \boldsymbol{v}_H\|_\Lambda \le cHk^{-1}\|\Lambda_h \boldsymbol{v}_h\|. \tag{6.2}$$

The proof of this result is completely analogous to the previous one.

Similarly, in Lemma 3.2, we need to replace the estimate

$$\|\operatorname{\mathbf{grad}}_h p\| + \|s\| \le cH\|\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}\|_{\boldsymbol{H}(\mathrm{div})}$$

by the estimate

$$\|\operatorname{\mathbf{grad}}_h p\| + \|s\| \le cHk^{-1}\|\boldsymbol{u} - \boldsymbol{P}_H \boldsymbol{u}\|_\Lambda. \tag{6.3}$$

Using (6.2), the proof of this result is completely analogous to the previous one.

Turning now to § 4 on domain decomposition, the proof of the second inequality in Theorem 4.1 carries over directly once k is introduced. The proof of the first inequality carries over directly until (4.5), for which the obvious replacement is

$$\sum_{j=1}^{J} \Lambda(\boldsymbol{v}_j, \boldsymbol{v}_j) \le c[(1 + k^2H^{-2})\|\boldsymbol{v} - \boldsymbol{v}_0\|^2 + k^2\|\operatorname{div}(\boldsymbol{v} - \boldsymbol{v}_0)\|^2].$$

When $H \le k$, we can complete the proof essentially as before using Lemma 3.1 to obtain from the above

$$\sum_{j=1}^{J} \Lambda(\boldsymbol{v}_j, \boldsymbol{v}_j) \le c[H^2 + k^2]\|\operatorname{div} \boldsymbol{v}\|^2 \le c\Lambda(\boldsymbol{v}, \boldsymbol{v}).$$

When $k < H$, we replace the $\boldsymbol{v}_0$ defined previously by $\boldsymbol{v}_0 = 0$. Then

$$\sum_{j=1}^{J} \Lambda(\boldsymbol{v}_j, \boldsymbol{v}_j) \le c[\|\boldsymbol{v} - \boldsymbol{v}_0\|^2 + k^2\|\operatorname{div}(\boldsymbol{v} - \boldsymbol{v}_0)\|^2] \le c\Lambda(\boldsymbol{v}, \boldsymbol{v}).$$

Turning to the section on multigrid, we need only check that the smoother still satisfies the conditions of Theorem 5.1. The proof of (5.1) is the same, requiring only the replacement of the $\|\cdot\|_{\boldsymbol{H}(\mathrm{div})}$ norm by the more general $\|\cdot\|_\Lambda$ norm. To

establish (5.2), we proceed as previously, obtaining instead of (5.6), the obvious replacement

$$\sum_{\nu\in\mathcal{N}_j} \Lambda(\bar{\boldsymbol{v}}_\nu, \bar{\boldsymbol{v}}_\nu) \le C(k^2 h_j^{-2}\|\bar{\boldsymbol{v}}\|^2 + \|\bar{\boldsymbol{v}}\|^2 + k^2\|\operatorname{div}\bar{\boldsymbol{v}}\|^2).$$

Applying (6.3), we get $\|\bar{\boldsymbol{v}}\| \le ch_j k^{-1}\|\boldsymbol{v}\|_\Lambda$. Since $\|\bar{\boldsymbol{v}}\| \le \|\boldsymbol{v}\|$ and $\|\operatorname{div}\bar{\boldsymbol{v}}\| = \|\operatorname{div}\boldsymbol{v}\|$, it follows immediately that

$$\sum_{\nu\in\mathcal{N}_j} \Lambda(\bar{\boldsymbol{v}}_\nu, \bar{\boldsymbol{v}}_\nu) \le C\|\boldsymbol{v}\|_\Lambda^2.$$

The remainder of the proof is unchanged.

7. Applications

In this section we give examples of how the preconditioners $\boldsymbol{\Theta}_h$ constructed in the previous sections can be used to develop efficient solution operators for the linear systems that arise from some finite element procedures.

To explain our approach, we first consider the abstract problem of finding x_h belonging to a normed finite dimensional vectorspace X_h and satisfying

$$\mathcal{A}_h x_h = f_h, \tag{7.1}$$

where $\mathcal{A}_h$ is a self-adjoint linear operator from X_h to its dual X_h^* and $f_h \in X_h^*$ is given. We think of $\mathcal{A}_h$ as coming from some discretization of a boundary value problem using finite elements of mesh size h and assume that the operator norms

$$\|\mathcal{A}_h\|_{\mathcal{L}(X_h,X_h^*)} \text{ and } \|\mathcal{A}_h^{-1}\|_{\mathcal{L}(X_h^*,X_h)} \text{ are bounded uniformly in } h. \tag{7.2}$$

In order to solve (7.1), we will use an iterative solution algorithm preconditioned by a positive-definite self-adjoint operator $\mathcal{B}_h : X_h^* \to X_h$. For example, we may use a preconditioned minimum residual iteration, or, if $\mathcal{A}_h$ is positive-definite, a preconditioned conjugate gradient iteration. Such an iterative scheme is efficient if the action of the $\mathcal{B}_h$ can be computed efficiently and if the magnitude of the eigenvalues of $\mathcal{B}_h\mathcal{A}_h$ can be bounded above and below by positive constants independent of h (this last property insures that the number of iterations needed to achieve a given factor of reduction of the error is bounded). Now, in light of (7.2), the desired eigenvalue bounds will follow if

$$\|\mathcal{B}_h\|_{\mathcal{L}(X_h^*,X_h)} \text{ and } \|\mathcal{B}_h^{-1}\|_{\mathcal{L}(X_h,X_h^*)} \text{ are bounded uniformly in } h. \tag{7.3}$$

Thus to efficiently solve (7.1), we simply require a computable positive-definite operator $\mathcal{B}_h$ for which (7.3) holds. We remark that the preconditioner $\mathcal{B}_h$ can be constructed without reference to the detailed structure of the operator $\mathcal{A}_h$, but depends only on the norm in X_h.

As a first example, consider an elliptic boundary value problem of the form

$$\operatorname{div}(a\,\mathbf{grad}\,p) = g \text{ in } \Omega, \qquad p = 0 \text{ on } \partial\Omega. \tag{7.4}$$

The data g is assumed to be $L^2(\Omega)$, while the coefficient matrix $a = \{a_{ij}(x)\}_{i,j=1}^2$ is assumed measurable, bounded, symmetric, and uniformly positive-definite on $\bar{\Omega}$.

Introducing the variable $\boldsymbol{u} = a\,\mathbf{grad}\,p$, we obtain the first order system

$$\boldsymbol{u} - a\,\mathbf{grad}\,p = 0 \text{ in } \Omega, \quad \operatorname{div}\boldsymbol{u} = g \text{ in } \Omega, \qquad p = 0 \text{ on } \partial\Omega. \tag{7.5}$$

A simple least squares approach characterizes $(\boldsymbol{u}, p)$ as the minimizer of the functional

$$J(\boldsymbol{v}, q) = \|\boldsymbol{v} - a\,\mathbf{grad}\,q\|_0^2 + \|\operatorname{div}\boldsymbol{v} - g\|_0^2$$

over $\boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1$. Defining $B : [\boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1] \times [\boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1] \to \mathbb{R}$ by

$$B(\boldsymbol{u}, p; \boldsymbol{v}, q) = (\boldsymbol{u} - a\,\mathbf{grad}\,p, \boldsymbol{v} - a\,\mathbf{grad}\,q) + (\operatorname{div}\boldsymbol{u}, \operatorname{div}\boldsymbol{v}),$$

it is easy to see that $(\boldsymbol{u}, p) \in \boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1$ is determined by the weak equations

$$B(\boldsymbol{u}, p; \boldsymbol{v}, q) = (g, \operatorname{div}\boldsymbol{v}) \text{ for all } (\boldsymbol{v}, q) \in \boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1.$$

In [25] (and, in greater generality, in [13]), it is shown that this bilinear form is symmetric and positive-definite, and so defines an inner product on $\boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1$ equivalent to the usual one.

By restricting the minimization to a finite dimensional subspace $X_h \subset \boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1$, we obtain an approximate solution, $(\boldsymbol{u}_h, p_h) \in X_h$. The convergence of this procedure was established for a wide variety of finite element spaces in [25]. Defining $\mathcal{A}_h : X_h \to X_h^*$ by

$$\langle \mathcal{A}_h x, y\rangle = B(x, y) \text{ for all } x, y \in X_h,$$

and $f_h \in X_h^*$ by

$$\langle f_h, (\boldsymbol{v}, q)\rangle = (g, \operatorname{div}\boldsymbol{v}),$$

the linear system determining the discrete solution can be written in the form (7.1). Since X_h is normed with the restriction of the norm in $\boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1$ and the norm in X_h^* is defined by duality, the bounds (7.2) follow directly from the equivalence of the B inner product with the inner product in $\boldsymbol{H}(\mathrm{div}) \times \mathring{H}^1$. Hence we need to construct a preconditioner $\mathcal{B}_h : X_h^* \to X_h$ for which (7.3) holds.

Now suppose that $X_h = \boldsymbol{V}_h \times W_h$ where $\boldsymbol{V}_h \subset \boldsymbol{H}(\mathrm{div})$ is a Raviart–Thomas space and W_h is some standard finite element subspace of $\mathring{H}^1$. As is usual we identify $\boldsymbol{V}_h^*$ with $\boldsymbol{V}_h$ so that

$$\|\boldsymbol{v}\|_{\boldsymbol{V}_h^*} = \sup_{\boldsymbol{w} \in \boldsymbol{V}_h} \frac{(\boldsymbol{v}, \boldsymbol{w})}{\|\boldsymbol{w}\|_{\boldsymbol{H}(\mathrm{div})}},$$

and similarly for W_h^*. The operators $\boldsymbol{\Theta}_h$ constructed in the previous sections map $\boldsymbol{V}_h^* \to \boldsymbol{V}_h$ and satisfy

$$\|\boldsymbol{\Theta}_h\|_{\mathcal{L}(\boldsymbol{V}_h^*, \boldsymbol{V}_h)} \text{ and } \|\boldsymbol{\Theta}_h^{-1}\|_{\mathcal{L}(\boldsymbol{V}_h, \boldsymbol{V}_h^*)} \text{ are bounded uniformly in } h.$$

Moreover, domain decomposition or multigrid can be used to construct $\Phi_h : W_h^* \to W_h$ such that

$$\|\Phi_h\|_{\mathcal{L}(W_h^*, W_h)} \text{ and } \|\Phi_h^{-1}\|_{\mathcal{L}(W_h, W_h^*)} \text{ are bounded uniformly in } h.$$

These are the natural requirements for a preconditioner for the Laplacian discretized in the usual way using the space W_h. Then, letting

$$\mathcal{B}_h = \begin{pmatrix} \boldsymbol{\Theta}_h & 0 \\ 0 & \Phi_h \end{pmatrix},$$

(7.2) follows directly.

To summarize this example: we may precondition the discrete least squares system using an $\boldsymbol{H}(\mathrm{div})$ preconditioner for the vector variable and a standard H^1 preconditioner for the scalar variable.

A second example is furnished by a mixed method solution to (7.4). This again proceeds from the first order system (7.5), but now the weak formulation is to find $(\boldsymbol{u}, p) \in \boldsymbol{H}(\mathrm{div}) \times L^2$ such that

$$\begin{aligned} (a^{-1}\boldsymbol{u}, \boldsymbol{v}) + (p, \mathrm{div}\, \boldsymbol{v}) &= 0 \text{ for all } \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}), \\ (\mathrm{div}\, \boldsymbol{u}, q) &= (g, q) \text{ for all } q \in L^2. \end{aligned} \tag{7.6}$$

We shall discretize this using the Raviart–Thomas space $\boldsymbol{V}_h$ of index r for $\boldsymbol{u}$ and the space S_h of discontinuous piecewise polynomials of degree r for p, so that the discrete weak formulation is to find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times S_h$ such that

$$\begin{aligned} (a^{-1}\boldsymbol{u}_h, \boldsymbol{v}) + (p_h, \mathrm{div}\, \boldsymbol{v}) &= 0 \text{ for all } \boldsymbol{v} \in \boldsymbol{V}_h, \\ (\mathrm{div}\, \boldsymbol{u}_h, q) &= (g, q) \text{ for all } q \in S_h. \end{aligned} \tag{7.7}$$

This is again a system of the form (7.1), where now $X_h = \boldsymbol{V}_h \times S_h$ and $\mathcal{A}_h : X_h \to X_h^*$ is self-adjoint but indefinite. The bounds (7.2) are a consequence of the stability of the Raviart–Thomas elements (cf., [10, Proposition II.1.3]).

Note that, since the norm on S_h is the L^2 norm, its dual norm coincides with itself. Thus the choice of preconditioner is obvious: we take

$$\mathcal{B}_h = \begin{pmatrix} \boldsymbol{\Theta}_h & 0 \\ 0 & I \end{pmatrix}, \tag{7.8}$$

where I is the identity on S_h, and then (7.3) holds. Vassilevski and Lazarov arrived at the same type of block diagonal preconditioner for the mixed system in [31], although they did not have available the simple multigrid and domain decomposition preconditioners for Λ_h which we constructed in §§ 4 and 5, and so suggested the more complicated operator from [33]. Let us comment on how this choice of preconditioner differs from other block diagonal preconditioners for the mixed method which have been considered. The coefficient operator of the continuous system (7.6),

$$\mathcal{A} = \begin{pmatrix} a^{-1} & -\mathbf{grad} \\ \mathrm{div} & 0 \end{pmatrix},$$

is an isomorphism from $\boldsymbol{H}(\mathrm{div}) \times L^2$ onto its dual $\boldsymbol{H}(\mathrm{div})^* \times L^2$, and the stability of the Raviart–Thomas discretization implied that similar mapping properties hold for the discrete operator $\mathcal{A}_h$. These mapping properties led naturally to our choice of preconditioner. However, it is also true that the continuous operator $\mathcal{A}$ defines an isomorphism from $\boldsymbol{L}^2 \times \mathring{H}^1$ onto its dual $\boldsymbol{L}^2 \times H^{-1}$. In fact, this is just a recasting of the standard $\mathring{H}^1 \to H^{-1}$ isomorphism for the Dirichlet problem for a second order elliptic equation. From a corresponding discrete result for the operator $\mathcal{A}_h$, and by arguing in an analogous manner as above, we are led to a preconditioner for the discrete operator $\mathcal{A}_h$ of the form

$$\tilde{\mathcal{B}}_h = \begin{pmatrix} \boldsymbol{I} & 0 \\ 0 & \Phi_h \end{pmatrix}. \tag{7.9}$$

where $\Phi_h : S_h \to S_h$ is a preconditioner for the discrete negative Laplace operator $-\Delta_h : S_h \to S_h$ given by $-\mathrm{div}\, \mathbf{grad}_h$. However, notice that S_h is not a subset of $\mathring{H}^1$ in the present case (indeed in the simplest case, S_h consists of piecewise constants), and the discrete Laplacian is not a standard, or even a local, operator, and so the definition of Φ_h is not obvious. From this point of view, this second approach seems less natural than the first. Despite this fact, most of the preconditioners for the

system (7.7) which are analyzed in the literature are of the form (7.9) or closely related to it. Such preconditioners are, for example, discussed in [27], [28], and [29]. In this context, we also mention that many authors have sought to avoid the solution of (7.7) when implementing mixed methods by using a reformulation of the mixed method as a positive-definite system. This can be accomplished by using the Schur complement as in algorithms of Uzawa type, via the introduction of additional Lagrange multipliers and elimination of the vector variable, or via the use of divergence free bases and elimination of the scalar variable. Examples of such approaches can be found in [4], [6], [8], [9], [14], [15],[16], [17], [20], [22], [24], [32].

Finally, in this section we shall consider an application of the preconditioning of the k-dependent operator associated to the bilinear form (6.1). Consider the system obtained by applying the mixed finite element method to the singular perturbation problem

$$k^2\,\Delta\, p - p = g \text{ in } \Omega, \qquad p = 0 \text{ on } \partial\Omega,$$

with $k \in (0, 1]$. A mixed formulation seeks $(\boldsymbol{u}, p) \in \boldsymbol{H}(\mathrm{div}) \times L^2$ such that

$$(\boldsymbol{u}, \boldsymbol{v}) + k(p, \operatorname{div} \boldsymbol{v}) = 0 \text{ for all } \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}),$$
$$k(\operatorname{div} \boldsymbol{u}, q) - (p, q) = (g, q) \text{ for all } q \in L^2.$$

The differential operator

$$\mathcal{A} = \begin{pmatrix} \boldsymbol{I} & -k\,\mathbf{grad} \\ k\operatorname{div} & -I \end{pmatrix}$$

defines an isomorphism from $\boldsymbol{H}(\mathrm{div}) \times L^2$ onto its dual. In fact, equipping $\boldsymbol{H}(\mathrm{div})$ with the norm

$$\boldsymbol{u} \mapsto (\|\boldsymbol{u}\|^2 + k^2 \|\operatorname{div} \boldsymbol{u}\|^2)^{1/2}$$

and choosing test functions $\boldsymbol{v} = \boldsymbol{u}$ and $q = k \operatorname{div} \boldsymbol{u} - p$, we easily see that the norms of $\mathcal{A}$ and $\mathcal{A}^{-1}$ are bounded uniformly with respect to k. Therefore we can precondition the mixed system with the block diagonal preconditioner (7.8), where $\boldsymbol{\Theta}_h$ is the domain decomposition or multigrid preconditioner for the k-dependent operator on $\boldsymbol{H}(\mathrm{div})$. In view of the uniformity of the bounds on $\boldsymbol{\Theta}_h$ discussed in the previous section, the resulting block diagonal preconditioner for the mixed method is effective uniformly with respect to k and h.

8. Numerical results

In this section we present numerical results which illustrate the multigrid convergence results of § 5 and their application to mixed methods, as discussed in § 7.

First we made a numerical study of the condition number of $\boldsymbol{\Lambda}_h : \boldsymbol{V}_h \to \boldsymbol{V}_h$ and the effect of preconditioning. We took the domain to be the unit square. Bisecting the square into two triangles by its negatively sloped diagonal yields the mesh of level 1, to which we associate the mesh size $h = 1$. The level m mesh has half the mesh size as the level $m-1$ mesh and is formed from it by subdividing each triangle into four similar triangles. Thus the level m mesh is a uniform triangulation of the square into 2^{2m-1} triangles and has mesh size $h = 1/2^{m-1}$. The space $\boldsymbol{V}_h$ is taken as Raviart–Thomas space of index 0 on this mesh.

TABLE 1. Condition numbers for the operator $\mathbf{\Lambda}_h$ and for the preconditioned operator $\mathbf{\Theta}_h\mathbf{\Lambda}_h$, and iteration counts to achieve an error reduction factor of 10^6.

level	h	elements	$\dim \boldsymbol{V}_h$	$\kappa(\mathbf{\Lambda}_h)$	$\kappa(\mathbf{\Theta}_h\mathbf{\Lambda}_h)$	iterations
1	1	2	5	38	1.00	1
2	1/2	8	16	153	1.32	4
3	1/4	32	56	646	1.68	6
4	1/8	128	208	2,650	2.17	6
5	1/16	512	800	10,670	2.34	8
6	1/32	2,048	3,136	42,810	2.40	8
7	1/64	8,192	12,416	–	–	8

Table 1 reports the spectral condition number of the discrete operator $\mathbf{\Lambda}_h$ and the preconditioned operator $\mathbf{\Theta}_h\mathbf{\Lambda}_h$ where $\mathbf{\Theta}_h$ is the V-cycle multigrid preconditioner of § 5 using one application of the additive Schwarz preconditioner (5.3) with the scaling factor η taken to be 1/2. (For the convergence theory of § 5 we assumed that $\eta \le 1/3$, but other approaches to the convergence theory can be used to raise this bound to 2/3, and we found the results slightly better with $\eta = 1/2$ than $\eta = 1/3$.) In order to determine the condition numbers, we computed the matrices corresponding to both the unpreconditioned and the preconditioned operator, and then calculated their largest and smallest eigenvalues. Of course, this is an expensive procedure which is never performed in a practical computation, but which we carried out to illustrate the theory. The fifth column of Table 1 clearly displays the expected growth of the condition number of $\mathbf{\Lambda}_h$ as $O(h^{-2})$, and the sixth column the boundedness of the condition number of the preconditioned operator $\mathbf{\Theta}_h\mathbf{\Lambda}_h$.

In addition to computing the condition numbers, we solved the equation (1.2) using the conjugate gradient method with $\mathbf{\Theta}_h$ as preconditioner. We arbitrarily took $\boldsymbol{f}$ to be the constant unit vectorfield in the vertical direction and started with an initial iterate of 0. The final column of Table 1 shows the number of iterations required to reduce the initial error by a factor of 10^6. As expected, the number of iterations appears to remain bounded as the mesh is refined.

As a second numerical study, we used the Raviart–Thomas mixed method to solve the factored Poisson equation

$$\boldsymbol{u} = \mathbf{grad}\, p, \quad \operatorname{div} \boldsymbol{u} = g \quad \text{in } \Omega, \qquad p = 0 \quad \text{on } \partial\Omega.$$

We chose $g = 2(x^2 + y^2 - x - y)$ so that $p = (x^2 - x)(y^2 - y)$. The discrete solution $(\boldsymbol{u}_h, p_h)$ belongs to the space $\boldsymbol{V}_h \times S_h$, with $\boldsymbol{V}_h$ the Raviart–Thomas space described above and S_h the space of piecewise constant functions on the same mesh. We solved the discrete equations both with a direct solver and by using the minimum residual method preconditioned with the block diagonal preconditioner having as diagonal blocks $\mathbf{\Theta}_h$ and the identity (as discussed in § 7). Full multigrid was used to initialize the minimum residual algorithm. That is, the computed solution at each level was used as an initial guess at the next finer level, beginning with the exact solution on the coarsest (two element) mesh. In Table 2 we show the condition number of the discrete operator $\mathcal{A}_h$ and of the preconditioned operator $\mathcal{B}_h\mathcal{A}_h$. While the former quantity grows linearly with h^{-1} (since this is a first order system), the latter remains small.

TABLE 2. Condition numbers for the indefinite operator $\mathcal{A}_h$ corresponding to the mixed system and for the preconditioned operator $\mathcal{B}_h\mathcal{A}_h$.

level	h	$\dim \boldsymbol{V}_h$	$\dim S_h$	$\kappa(\mathcal{A}_h)$	$\kappa(\mathcal{B}_h\mathcal{A}_h)$
1	1	5	2	8.25	1.04
2	1/2	16	8	15.0	1.32
3	1/4	56	32	29.7	1.68
4	1/8	208	128	59.6	2.18
5	1/16	800	512	119	2.34

TABLE 3. Percent relative L^2 errors for the mixed method with three different solvers for the linear system: direct, 4 iterations of the minimum residual method, and 8 iterations of minimum residual.

				direct solve		minres 4 iter.		minres 8 iter.	
level	h	$\dim \boldsymbol{V}_h$	$\dim S_h$	$\boldsymbol{u}$	p	$\boldsymbol{u}$	p	$\boldsymbol{u}$	p
1	1	5	2	33.33	33.33	33.33	33.33	33.33	33.33
2	1/2	16	8	38.90	7.49	38.90	7.46	38.90	7.49
3	1/4	56	32	23.44	2.89	23.50	9.02	23.44	2.89
4	1/8	208	128	12.30	0.84	12.38	4.48	12.30	0.90
5	1/16	800	512	6.22	0.22	6.26	1.92	6.22	0.24
6	1/32	3,136	2,048	3.12	0.05	3.14	0.75	3.12	0.06
7	1/64	12,416	8,192	1.56	0.01	1.57	0.32	1.56	0.02

Finally, we studied the contribution of the preconditioned minimum residual method to the solution error. To measure the error in the vector variable $\boldsymbol{u}$, we used the relative L^2 error $\|\boldsymbol{u} - \boldsymbol{u}_c\|_0/\|\boldsymbol{u}\|_0$, where $\boldsymbol{u}_c$ represents the computed solution. The L^2 norms were computed using the three point quadrature rule (with edge midpoints as quadrature points) on each element and this is reported in Table 3 as a percent. Note that the reported error involves both the discretization error of the mixed method and further errors introduced by the linear solution process. To measure the error in the scalar variable p, we compared the piecewise constant computed solution to the piecewise constant function p^* whose value is obtained on each element by averaging the values of the exact solution at the three quadrature points of the element; that is, we report $\|p^* - p_h\|_0/\|p^*\|_0$ as a percent. We do this because p^* is a superconvergent quantity: $\|p^* - p_h\|_0 = O(h^2)$. We see from Table 3 that for the vector variable $\boldsymbol{u}$, the full accuracy of the approximation is achieved with only four iterations of the minimum residual method, even when the system has over 20,000 unknowns. To maintain the full accuracy of the superconvergent approximation to p^*, more iterations are needed, but for all practical purposes 8 iterations are sufficient, even for the finest mesh.

APPENDIX A

In the foregoing analysis we have assumed that the polygonal domain Ω is convex. In this appendix we show that the results of § 4 on domain decomposition can be proven without assuming convexity. A careful examination of that section reveals

that the only time convexity used was in the application of (3.9) to bound the first summand on the right-hand side of (4.5). However, the proof of (3.9) depended on 2-regularity of the Laplacian, and so fails for a non-convex polygonal domain.

In the proof of the second inequality in Theorem 4.1 one did not use convexity, so only the proof of the first inequality needs to be adapted to the non-convex case. Recall that the key to establishing the first inequality was to show that we can decompose any $\boldsymbol{u} \in \boldsymbol{V}_h$ as $\sum_{j=0}^{J} \boldsymbol{u}_j$ with $\boldsymbol{u}_j \in \boldsymbol{V}_j$ and satisfying (4.2). To do so, we split $\boldsymbol{u}$ as $\mathbf{grad}_h\, p + \mathbf{curl}\, w$ via the discrete Helmholtz decomposition, and set $\boldsymbol{v} = \mathbf{grad}_h\, p$. In proving (4.2) in the convex case, we defined an approximation $(\boldsymbol{v}_0, p_0) \in \boldsymbol{V}_0 \times S_0$ by $\boldsymbol{v}_0 = \mathbf{grad}_0\, p_0$, $\operatorname{div} \boldsymbol{v}_0 = Q_0 \operatorname{div} \boldsymbol{v}$, where Q_0 is the L^2 projection into S_0, and then we bounded the difference $\|\boldsymbol{v} - \boldsymbol{v}_0\|$ by $cH\|\operatorname{div} \boldsymbol{v}\|$ using (3.9). Since this fails in the non-convex case, we now introduce an alternate approximation, namely we define $(\tilde{\boldsymbol{v}}_h, \tilde{p}_h) \in \boldsymbol{V}_h \times S_h$ by $\tilde{\boldsymbol{v}}_h = \mathbf{grad}_h\, \tilde{p}_h$, $\operatorname{div} \tilde{\boldsymbol{v}}_h = Q_0 \operatorname{div} \boldsymbol{v}$. We shall show below that

$$\|\boldsymbol{v} - \tilde{\boldsymbol{v}}_h\| \le cH\|\operatorname{div} \boldsymbol{v}\|. \tag{A.1}$$

Assuming (A.1) for the moment, we now complete the proof of Theorem 4.1 without the assumption of convexity. From the definitions it follows that $\operatorname{div}(\tilde{\boldsymbol{v}}_h - \boldsymbol{v}_0) = 0$ and hence that $\tilde{\boldsymbol{v}}_h - \boldsymbol{v}_0 = \mathbf{curl}\, \rho$ for some $\rho \in W_h$ of mean value zero. Thus,

$$\boldsymbol{u} = (\boldsymbol{v} - \tilde{\boldsymbol{v}}_h) + (\tilde{\boldsymbol{v}}_h - \boldsymbol{v}_0) + \boldsymbol{v}_0 + \mathbf{curl}\, w = (\boldsymbol{v} - \tilde{\boldsymbol{v}}_h + \boldsymbol{v}_0) + \mathbf{curl}(w + \rho).$$

Note that $\|w\|_1 \le c\|\boldsymbol{u}\|_{\boldsymbol{H}(\operatorname{div})}$ and $\|\rho\|_1 \le c(\|\tilde{\boldsymbol{v}}_h\| + \|\boldsymbol{v}_0\|) \le c\|\operatorname{div} \boldsymbol{v}\| \le c\|\boldsymbol{u}\|_{\boldsymbol{H}(\operatorname{div})}$. We are now ready to choose the $\boldsymbol{u}_j$. First, we decompose $w + \rho$ as $\sum_{j=0}^{J} w_j$ with $w_j \in W_j$ and $\sum_{j=0}^{J} \|w_j\|_1^2 \le c\|w + \rho\|_1^2$. Then $\mathbf{curl}(w + \rho) = \sum_{j=0}^{J} \mathbf{curl}\, w_j$, $\mathbf{curl}\, w_j \in \boldsymbol{V}_j$, and

$$\sum_{j=0}^{J} \Lambda(\mathbf{curl}\, w_j, \mathbf{curl}\, w_j) \le c\|w + \rho\|_1^2 \le c\Lambda(\boldsymbol{u}, \boldsymbol{u}).$$

It thus remains to decompose $\boldsymbol{v} - \tilde{\boldsymbol{v}}_h + \boldsymbol{v}_0$ as $\sum_{j=0}^{J} \boldsymbol{v}_j$ with $\boldsymbol{v}_j \in \boldsymbol{V}_j$ such that $\sum_{j=0}^{J} \Lambda(\boldsymbol{v}_j, \boldsymbol{v}_j) \le c\Lambda(\boldsymbol{v}, \boldsymbol{v})$. To do so, we proceed in an analogous manner to the earlier proof, setting $\boldsymbol{v}_j = \boldsymbol{\Pi}_h[\theta_j(\boldsymbol{v} - \tilde{\boldsymbol{v}}_h)]$ for $j \ge 1$. This leads to the inequality

$$\sum_{j=1}^{J} \Lambda(\boldsymbol{v}_j, \boldsymbol{v}_j) \le c[(1 + H^{-2})\|\boldsymbol{v} - \tilde{\boldsymbol{v}}_h\|^2 + \|\operatorname{div}(\boldsymbol{v} - \tilde{\boldsymbol{v}}_h)\|^2],$$

which is analogous to (4.5), and the proof is completed by invoking (A.1) and arguing as before.

It thus remains to prove (A.1). From the definitions of $\boldsymbol{v}$ and $\tilde{\boldsymbol{v}}_h$, we get that

$$\begin{aligned} &(\boldsymbol{v} - \tilde{\boldsymbol{v}}_h, \boldsymbol{w}) + (p - \tilde{p}_h, \operatorname{div} \boldsymbol{w}) = 0, \text{ for all } \boldsymbol{w} \in \boldsymbol{V}_h, \\ &(\operatorname{div}[\boldsymbol{v} - \tilde{\boldsymbol{v}}_h], q) = ([I - Q_0] \operatorname{div} \boldsymbol{v}, q), \text{ for all } q \in S_h. \end{aligned}$$

Choosing $\boldsymbol{w} = \boldsymbol{v} - \tilde{\boldsymbol{v}}_h$ and $q = p - \tilde{p}_h$ and subtracting the equations, we get

$$\begin{aligned} \|\boldsymbol{v} - \tilde{\boldsymbol{v}}_h\|^2 &= -([I - Q_0] \operatorname{div} \boldsymbol{v}, p - \tilde{p}_h) = -(\operatorname{div} \boldsymbol{v}, [I - Q_0][p - \tilde{p}_h]) \\ &\le \|\operatorname{div} \boldsymbol{v}\| \|(I - Q_0)(p - \tilde{p}_h)\|. \end{aligned}$$

Hence (A.1) will follow if we can show that

$$\|(I - Q_0)(p - \tilde{p}_h)\| \le CH\|\operatorname{\mathbf{grad}}_h(p - \tilde{p}_h)\| = CH\|\boldsymbol{v} - \tilde{\boldsymbol{v}}_h\|.$$

This is an immediate consequence of the following lemma.

Lemma A.1. *There exists a constant c independent of h and H such that*

$$\|p - Q_H p\| \le cH\|\operatorname{\mathbf{grad}}_h p\|, \text{ for all } p \in S_h.$$

Proof. We first note that this result was established in the case of a convex domain in Lemma 3.1. To extend the proof to the nonconvex case, we first consider a single triangle $T = T_H$ in the coarse mesh composed of triangles $t = t_h$ in the fine mesh. We define the spaces

$$\boldsymbol{V}_h^T = \{\, \boldsymbol{v} \in \boldsymbol{V}_h : \boldsymbol{v} = 0 \text{ on } \Omega \setminus T \,\}, \qquad S_h^T = \{p \in S_h : p = 0 \text{ on } \Omega \setminus T\,\},$$

and an operator $\operatorname{\mathbf{grad}}_h^T : S_h^T \to \boldsymbol{V}_h^T$ defined for $p \in S_h^T$ by

$$(\operatorname{\mathbf{grad}}_h^T p, \boldsymbol{v}) = -(p, \operatorname{div} \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in \boldsymbol{V}_h^T.$$

Now, define $p_T \in S_h^T$ by $p_T = p$ on T and $p_T = 0$ on $\Omega \setminus T$. Then

$$\begin{aligned}\|\operatorname{\mathbf{grad}}_h^T p_T\|_{0,T}^2 &= (\operatorname{\mathbf{grad}}_h^T p_T, \operatorname{\mathbf{grad}}_h^T p_T) = -(p_T, \operatorname{div}\operatorname{\mathbf{grad}}_h^T p_T)\\ &= -(p, \operatorname{div}\operatorname{\mathbf{grad}}_h^T p_T) = (\operatorname{\mathbf{grad}}_h p, \operatorname{\mathbf{grad}}_h^T p_T)\\ &= (\operatorname{\mathbf{grad}}_h p, \operatorname{\mathbf{grad}}_h^T p_T)_T \le \|\operatorname{\mathbf{grad}}_h p\|_{0,T}\|\operatorname{\mathbf{grad}}_h^T p_T\|_{0,T},\end{aligned}$$

so $\|\operatorname{\mathbf{grad}}_h^T p_T\|_{0,T} \le \|\operatorname{\mathbf{grad}}_h p\|_{0,T}$. Now since the triangle T is convex, we may apply (3.8) to p_T, obtaining $\|p_T - Q_H p_T\|_{0,T} \le cH\|\operatorname{\mathbf{grad}}_h^T p_T\|_{0,T}$. Therefore

$$\begin{aligned}\|p - Q_H p\|^2 &= \sum_T \|p_T - Q_H p_T\|_{0,T}^2 \le c^2H^2 \sum_T \|\operatorname{\mathbf{grad}}_h^T p_T\|_{0,T}^2\\ &\le c^2H^2 \sum_T \|\operatorname{\mathbf{grad}}_h p\|_{0,T}^2 \le c^2H^2\|\operatorname{\mathbf{grad}}_h p\|^2. \quad \square\end{aligned}$$

Appendix B

We now give a proof of the abstract convergence result, Theorem 5.1, for the V-cycle, and the identity (2.1) for the additive Schwarz operator.

Proof of Theorem 5.1. We shall prove by induction on i that

$$0 \le \Lambda([I - \Theta_i\Lambda_i]u, u) \le \delta\Lambda(u, u) \text{ for all } u \in V_j. \tag{B.1}$$

The result for $i = 1$ is obvious since $\Theta_1 = \Lambda_1^{-1}$. Now assume that (B.1) holds for $i = j - 1$. Setting $K_j = I - R_j\Lambda_j$, it is straightforward to derive the recurrence relation (cf. [5])

$$I - \Theta_j\Lambda_j = K_j^m[(I - P_{j-1}) + (I - \Theta_{j-1}\Lambda_{j-1})P_{j-1}]K_j^m.$$

The lower bound easily follows from this identity and the inductive hypothesis. For the upper bound, we use the induction hypothesis to obtain

$$\begin{aligned}\Lambda([I - \Theta_j\Lambda_j]u, u) &\le \Lambda([I - P_{j-1}]K_j^m u, K_j^m u) + \delta\Lambda(P_{j-1}K_j^m u, K_j^m u)\\ &= (1 - \delta)\Lambda([I - P_{j-1}]K_j^m u, K_j^m u) + \delta\Lambda(K_j^m u, K_j^m u).\end{aligned}$$

Now

$$\begin{aligned}\Lambda([I-P_{j-1}]K_j^m u,[I-P_{j-1}]K_j^m u) &= \Lambda([I-P_{j-1}]K_j^m u,K_j^m u)\\ &= ([I-P_{j-1}]K_j^m u,\Lambda_j K_j^m u)\\ &= (R_j^{-1}[I-P_{j-1}]K_j^m u,R_j\Lambda_j K_j^m u)\\ &\le (R_j^{-1}[I-P_{j-1}]K_j^m u,[I-P_{j-1}]K_j^m u)^{1/2}(R_j\Lambda_j K_j^m u,\Lambda_j K_j^m u)^{1/2}\\ &\le \sqrt{\alpha}\Lambda([I-P_{j-1}]K_j^m u,[I-P_{j-1}]K_j^m u)^{1/2}(R_j\Lambda_j K_j^m u,\Lambda_j K_j^m u)^{1/2}.\end{aligned}$$

Hence,

$$\Lambda([I-P_{j-1}]K_j^m u,K_j^m u)\le\alpha(R_j\Lambda_j K_j^m u,\Lambda_j K_j^m u)=\alpha\Lambda([I-K_j]K_j^{2m}u,u).$$

It follows from the positive semidefiniteness of R_j and from (5.1) that the spectrum of K_j is contained in the interval $[0,1]$. Therefore $\Lambda([I-K_j]K_j^{2m}u,u)\le\Lambda([I-K_j]K_j^i u,u)$ for $i\le 2m$, whence

$$\Lambda([I-K_j]K_j^{2m}u,u)\le\frac{1}{2m}\sum_{i=0}^{2m-1}\Lambda([I-K_j]K_j^i u,u)=\frac{1}{2m}\Lambda([I-K_j^{2m}]u,u).$$

Combining these results, we obtain

$$\begin{aligned}\Lambda([I-\Theta_j\Lambda_j]u,u) &\le (1-\delta)\frac{\alpha}{2m}\Lambda([I-K_j^{2m}]u,u)+\delta\Lambda(K_j^m u,K_j^m u)\\ &= (1-\delta)\frac{\alpha}{2m}\Lambda(u,u)+\left[\delta-(1-\delta)\frac{\alpha}{2m}\right]\Lambda(K_j^m u,K_j^m u).\end{aligned}$$

The result now follows by choosing

$$\delta=(1-\delta)\frac{\alpha}{2m},\quad\text{i.e.,}\quad\delta=\frac{\alpha}{\alpha+2m}.\quad\square$$

Proof of (2.1) . Recalling that $\Theta=\sum_j P_j B^{-1}$, and writing $v=\sum_j v_j$, we have

$$\begin{aligned}(\Theta^{-1}v,v) &= \sum(\Theta^{-1}v,v_j)=\sum(BB^{-1}\Theta^{-1}v,v_j)=\sum(BP_jB^{-1}\Theta^{-1}v,v_j)\\ &\le \sum(BP_jB^{-1}\Theta^{-1}v,P_jB^{-1}\Theta^{-1}v)^{1/2}(Bv_j,v_j)^{1/2}\\ &\le \left[\sum(BP_jB^{-1}\Theta^{-1}v,B^{-1}\Theta^{-1}v)\right]^{1/2}\left[\sum(Bv_j,v_j)\right]^{1/2}\\ &= \left[(Bv,B^{-1}\Theta^{-1}v)\right]^{1/2}\left[\sum(Bv_j,v_j)\right]^{1/2}\\ &\le \left[(v,\Theta^{-1}v)\right]^{1/2}\left[\sum(Bv_j,v_j)\right]^{1/2}.\end{aligned}$$

Hence, $(\Theta^{-1}v,v)\le\sum(Bv_j,v_j)$, and since the splitting was arbitrary,

$$(\Theta^{-1}v,v)\le\inf\sum(Bv_j,v_j),$$

where the infimum is over all decompositions $v=\sum_j v_j$. For the choice $v_j=P_jB^{-1}\Theta^{-1}v$, we have that $(\Theta^{-1}v,v)=\sum_j(Bv_j,v_j)$, and so (2.1) holds. $\square$

Acknowledgements

The authors are grateful to Professors J. Bramble, J. Pasciak, and J. Xu for many useful discussions.

References

1. D. N. Arnold, R. S. Falk, and R. Winther, *Preconditioning discrete approximations of the Reissner-Mindlin plate model*, Preprint (1995).
2. D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the V-cycle*, SIAM J. Numer. Anal. **20** (1983), 967-975. MR **85h:**65233
3. J. H. Bramble, *Multigrid methods*, Pitman Research Notes in Mathematics Series **294** (1993). MR **95b:**65002
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), 1-17. MR **89m:**65097a
5. J. H. Bramble and J. E. Pasciak, *New estimates for multilevel algorithms including the V-cycle*, Math. Comp. **60** (1993), 447-471. MR **94a:**65064
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. **34** (1997).
7. J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp. **57** (1991), 1-21. MR **92d:**65094
8. J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms with nonnested spaces and noninherited quadratic forms*, Math. Comp. **56** (1991), 1-34. MR **91h:**65159
9. S. C. Brenner, *A multigrid algorithm for the lowest-order Raviart-Thomas mixed triangular finite element method*, SIAM J. Numer. Anal. **29** (1992), 647-678. MR **93j:**65175
10. F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, vol. 15, Springer Verlag (1991). MR **92d:**65187
11. 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), 125-151. MR **92e:**73030
12. Z. Cai, C. I. Goldstein, J. E. Pasciak, *Multilevel iteration for mixed finite element systems with penalty*, SIAM J. Sci. Comput. **14** (1993), 1072–1088. MR **94h:**65116
13. Z. Cai, R. Lazarov, T. Manteuffel, and S. McCormick, *First-order system least squares for second-order partial differential equations: Part I*, SIAM J. Numer. Anal. **31** (1994), 1785-1799. MR **95i:**65133
14. Z. Chen, *Equivalence between and multigrid algorithms for mixed and nonconforming methods for second order elliptic problems*, East-West J. Numer. Math. **4** (1996). CMP 96:13
15. Z. Chen, R. E. Ewing, and R. Lazarov, *Domain decomposition algorithms for mixed methods for second order elliptic problems*, Math. Comp. **65** (1996), 467-490. MR **96g:**65117
16. L. C. Cowsar, *Dual-variable Schwarz methods for mixed finite elements*, Report TR93–09, Rice University (1993).
17. L. C. Cowsar, J. Mandel, and M. F. Wheeler, *Balancing domain decomposition for mixed finite elements*, Math. Comp. **64** (1995), 989-1016. MR **95j:**65161
18. M. Dryja and O. B. Widlund, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, Houston, Texas 1989, T. Chan et. al., eds., SIAM Publications, Philadelphia, 1990, pp. 3-21. MR **91m:**65294
19. M. Dryja and O. B. Widlund, *Some domain decomposition algorithms for elliptic problems*, Iterative methods for large linear systems, L. Hayes and D. Kincaid, eds., Academic Press, San Diego, 1990, pp. 273-291. CMP 90:07
20. H. C. Elman and G. Golub, *Inexact and preconditioned Uzawa algorithms for saddle point problems*, SIAM J. Numer. Anal. **31** (1994), 1645-1661. MR **95f:**65065
21. R. S. Falk and J. E. Osborn, *Error estimates for mixed methods*, R.A.I.R.O. Numerical Analysis **14** (1980), 249-277. MR **82j:**65076

22. R. Glowinski and M. Wheeler, *Domain decomposition and mixed finite element methods for elliptic problems*, First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski et. al., eds., SIAM Publications, Philadelphia, 1988, pp. 144-172. MR **90a:**65237
23. P. Lin, *A sequential regularization method for time-dependent incompressible Navier–Stokes equations*, to appear in SIAM J. Numer. Anal. **34** (1997).
24. T. P. Mathew, *Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part I: algorithms and numerical results*, Numer. Math. **65** (1993), 445-468. MR **94m:**65171
25. A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM J. Numer. Anal. **31** (1994), 1368-1377. MR **95f:**65206
26. P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2-order elliptic problems*, Mathematical Aspects of Finite Element Methods, Lecture Notes in Mathematics **606** (1977), 295-315. MR **58:**3547
27. T. Rusten and R. Winther, *A preconditioned iterative method for saddle point problems*, SIAM J. Matrix Anal. Appl. **13** (1992), 887-904. MR **93a:**65043
28. T. Rusten and R. Winther, *Substructure preconditioners for ellptic saddle point problems*, Math. Comp. **60** (1993), 23-48. MR **93d:**65108
29. T. Rusten, P. S. Vassilevski, and R. Winther, *Interior penalty preconditioners for mixed finite element approximations of elliptic problems*, Math. Comp. **65** (1996), 447-466. MR **96j:**65127
30. T. Rusten, P. S. Vassilevski, and R. Winther, *Domain embedding preconditioners for mixed systems*, in preparation.
31. P. Vassilevski and R. Lazarov, *Preconditioning saddle-point problems arising from mixed finite element discretizations of elliptic equations*, to appear in Numer. Lin. Algebra Appl.
32. P. S. Vassilevski and J. Wang, *An application of the abstract multilevel theory to nonconforming finite element methods*, SIAM J. Num. Anal. **32** (1995), 235–248. MR **96a:**65173
33. P. S. Vassilevski and J. Wang, *Multilevel iterative methods for mixed finite element discretizations of elliptic problems*, Numer. Math. **63** (1992), 503-520. MR **93j:**65187
34. J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev. **34** (1992), 581-613. MR **93k:**65029

Department of Mathematics, The Pennsylvania State University, University Park, Pennsylvania 16802
E-mail address: dna@math.psu.edu

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903
E-mail address: falk@math.rutgers.edu

Department of Informatics, University of Oslo, Oslo, Norway
E-mail address: ragnar@ifi.uio.no