INTERIOR PENALTY PRECONDITIONERS
FOR MIXED FINITE ELEMENT APPROXIMATIONS
OF ELLIPTIC PROBLEMS

TORGEIR RUSTEN, PANAYOT S. VASSILEVSKI, AND RAGNAR WINTHER

ABSTRACT. It is established that an interior penalty method applied to second-order elliptic problems gives rise to a local operator which is spectrally equivalent to the corresponding nonlocal operator arising from the mixed finite element method. This relation can be utilized in order to construct preconditioners for the discrete mixed system. As an example, a family of additive Schwarz preconditioners for these systems is constructed. Numerical examples which confirm the theoretical results are also presented.

1. INTRODUCTION

The purpose of this paper is to discuss the construction of preconditioners for the discrete problems arising from mixed finite element discretizations of second-order elliptic boundary value problems. In particular, we are interested in preconditioners constructed by domain decomposition. The main observation made in this paper is that, under suitable assumptions, there is a spectral equivalence between the local operator which arises from the interior penalty method studied by Arnold [1] and the corresponding nonlocal operator which is generated by the mixed method. Hence, any reasonable preconditioner for the interior penalty operator is also a suitable preconditioner for the corresponding mixed system. As an example of this approach we will use the interior penalty method to generate additive Schwarz preconditioners for the mixed system.

Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain with boundary $\partial\Omega$. For a given $f \in L^2(\Omega)$ and a given measurable, symmetric, uniformly positive definite coefficient matrix $k = \{k_{i,j}(x)\}_{i,j=1}^2$ on $\bar{\Omega}$ we consider the second-order elliptic problem

$$(1.1) \qquad \begin{aligned} -\nabla \cdot k\nabla p &= f \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

If the equation (1.1) is discretized by a conforming finite element method, a symmetric, positive definite discrete system is obtained. In order to design effective iterative methods for these discrete elliptic systems, it is necessary to construct

suitable preconditioners. If machines with a parallel architecture are to be used, then domain decomposition preconditioners have proved to be very effective. These preconditioners are constructed by utilizing decompositions of the solution space into subspaces corresponding to a covering of Ω by a collection of subdomains. If the subdomains are nonoverlapping, such that the subdomains define a coarse grid on Ω, the complete preconditioner is defined from a composition of local problems on each subdomain, proper continuity equations on the interior boundaries and a suitable coarse global problem. For the discussion of nonoverlapping domain decomposition methods for conforming finite element discretizations we refer to Bjørstad and Widlund [6] and Bramble, Pasciak and Schatz [10, 9] and the recent survey papers by Dryja, Smith and Widlund [20] and Xu [38]. The advantage of overlapping domain decomposition methods, or Schwarz methods, is partly that there is no need for continuity equations on the interior boundaries. Also, these methods are more robust with respect to the choice of local solvers. Schwarz methods for conforming finite element approximations are discussed by Lions [26], Dryja and Widlund [19], Matsokin and Nepomnyaschikh [27], Dryja, Smith and Widlund [20] and Xu [38].

The mixed finite element method for (1.1) can be derived from a system formulation of the problem. By introducing the "velocity" $\underline{u} = -k\nabla p$ as a new variable, we obtain the system

$$k^{-1}\underline{u} + \nabla p = 0,$$
$$\nabla \cdot \underline{u} = f.$$

When this system is discretized, using piecewise polynomial finite element spaces with basis functions of local support, we obtain a linear system with a symmetric, indefinite coefficient matrix of the form

$$(1.2) \qquad \begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix}.$$

Here, A is weighted mass matrix, defined from k^{-1}, while B and B^* are matrix representations of discrete analogs of the negative divergence and the gradient operators, respectively. In order to speed up the convergence of an iterative method for (1.2), it is necessary to construct a preconditioner for the positive definite matrix $L \equiv BA^{-1}B^*$, i.e. the Schur complement (cf. §3 below). This operator is a discrete analog of the second-order differential operator $-\nabla \cdot (k\nabla)$ which defines (1.1). However, a major difficulty with the approximation of this operator introduced by the mixed method is that it is in general *not* a local operator, owing to the appearance of the inverse of the mass matrix A. Also, if k is a nondiagonal matrix, the matrix A cannot be easily inverted. In fact, even if k is diagonal, the matrix A can only be easily inverted if special finite element spaces like the rectangular Raviart–Thomas elements are chosen (cf. Ewing and Wheeler [22]).

The fact that the mixed method introduces a nonlocal approximation of the operator $-\nabla \cdot (k\nabla)$ has consequences for the design of preconditioners for the mixed system. In particular, this is the case for a domain decomposition approach, since certain discrete problems on the subdomains will be nonlocal. One way to overcome this difficulty is to extend the discrete spaces where approximations of the variables $\underline{u}$ and p are sought in order to localize the operator L. This is for example done in the hybrid version of the mixed finite element method, where the continuity requirements on the discrete "velocity space" are relaxed at the expense of

introducing Langrange multipliers on all the element edges. For a general discussion of the hybrid mixed finite element method we refer to Brezzi and Fortin [15]. The development of domain decomposition preconditioners for the hybrid method is discussed by Glowinski and Wheeler [25], Cowsar [16] and Cowsar, Mandel and Wheeler [17].

In Rusten and Winther [32] an alternative to the standard hybrid version of the mixed method is discussed, where the continuity requirements are only relaxed on the edges of a coarse grid. This relaxation has the effect that the operator L becomes local with respect to the elements of the coarse grid. In [32] this property is used in order to define a nonoverlapping domain decomposition preconditioner for the mixed finite element method.

In contrast to the approaches described above we shall in this paper *not* work with any hybridization of the mixed finite element method. Instead of partly changing the method in order to localize the operator L, we shall establish that this operator is spectrally equivalent to the local operator which arises when the interior penalty method of [1] is applied to the problem (1.1). Hence, the interior penalty method can be used to construct preconditioners for L. Motivated by this fact, we shall analyze a family of additive Schwarz preconditioners for the interior penalty method. In this way we indirectly also develop a theory for additive Schwarz preconditioners for the discrete systems obtained from the mixed finite element method.

The difference between our approach and the one taken in Cowsar [16] (cf., in particular §7 of [16]) is that we do not relate the mixed approximation to a conforming approximation. We should also mention that efficient multigrid preconditioners for the mixed systems can be derived using the techniques developed by Bramble, Pasciak and Xu [12] (cf. also Vassilevski and Wang [35, 36]).

In §2 we give a brief review of the mixed finite element method and state the main assumptions that will be needed in later sections. A discussion of iterative methods and preconditioning for discrete saddle point problems is given in §3. In this section we also derive the spectral equivalence between the mixed method and the interior penalty method. The results for the additive Schwarz preconditioner are derived in §4. Finally, in §5 we present some numerical experiments.

2. PRELIMINARIES

For any function space X the associated norm will be denoted by $\| \cdot \|_X$. If $H^m(\Omega)$ denotes the L^2-based Sobolev space of order m on Ω, we introduce the simpler notation $\| \cdot \|_m$ instead of $\| \cdot \|_{H^m(\Omega)}$. Furthermore, $\| \cdot \|_{m,\Omega_1}$ will be used to denote $\| \cdot \|_{H^m(\Omega_1)}$ when $\Omega_1 \neq \Omega$.

In order to introduce the mixed formulation of the problem (1.1), the space $H(\operatorname{div}; \Omega) \equiv \left\{ \underline{\chi} \in L^2(\Omega)^2 : \ \nabla \cdot \underline{\chi} \in L^2(\Omega) \right\}$ is required. The norm on this space is defined by

$$\|\underline{\chi}\|_{H(\operatorname{div};\Omega)} = \left(\|\nabla \cdot \underline{\chi}\|_0^2 + \|\underline{\chi}\|_0^2 \right)^{\frac{1}{2}}.$$

Using integration by parts and having in mind the boundary condition $p = 0$ on $\partial\Omega$, one obtains the following mixed formulation of (1.1):

Find $\underline{u} \in H(\operatorname{div}; \Omega)$ and $p \in L^2(\Omega)$ such that

$$(2.1) \qquad \begin{aligned} a(\underline{u}, \underline{\chi}) + b(\underline{\chi}, p) &= 0 & &\text{for all } \underline{\chi} \in H(\operatorname{div}; \Omega), \\ b(\underline{u}, w) &= -(f, w) & &\text{for all } w \in L^2(\Omega). \end{aligned}$$

Here,

$$a(\underline{\chi}, \underline{\varphi}) = \int_\Omega k^{-1} \underline{\chi} \cdot \underline{\varphi} \, dx,$$

$$b(\underline{\chi}, w) = -\int_\Omega w \nabla \cdot \underline{\chi} \, dx,$$

$$(f, g) = \int_\Omega fg \, dx.$$

We will also use the standard inner product $(\underline{\chi}, \underline{\varphi})$ for functions $\underline{\chi}$ and $\underline{\varphi} \in L^2(\Omega)^2$.

We consider a quasi-uniform family of triangulation $\mathcal{T} = \mathcal{T}_h$ of Ω into triangles or rectangles, where $h \in (0, 1]$ is a parameter which indicates the grid size, i.e., h is a characteristic diameter of the elements. An adequate approximation of the problem (2.1) can then be obtained using piecewise polynomial finite element spaces $\mathbf{V} = \mathbf{V}_h \subset H(\text{div}; \Omega)$ and $W = W_h \subset L^2(\Omega)$ with respect to the triangulation $\mathcal{T}_h$. The space of piecewise constants with respect to $\mathcal{T}_h$ is denoted W_h^c. We assume that $W_h^c \subset W_h$. Furthermore, throughout this paper the spaces $\mathbf{V}_h$ and W_h are required to satisfy the Babuška–Brezzi condition, i.e., there is a positive constant β_0, independent of the mesh parameter h, such that

$$(2.2) \qquad \beta_0 \|w\|_{L^2(\Omega)} \leq \sup_{\underline{\chi} \in \mathbf{V}_h} \frac{b(\underline{\chi}, w)}{\|\underline{\chi}\|_{H(\text{div};\Omega)}} \quad \text{for all } w \in W_h.$$

In fact, we shall below state two assumptions, (A-I) and (A-II), which will, in particular, imply (2.2). Spaces that satisfy the stability condition (2.2) are proposed by Raviart–Thomas [30] for triangular and rectangular elements, in Brezzi, Douglas, Fortin and Marini [13], Brezzi, Douglas and Marini [14], Douglas and Wang [18] and can be also found in the texts of Girault and Raviart [24] and Brezzi and Fortin [15].

The finite element discretization of (2.1) then reads as follows: Find $\underline{u}_h \in \mathbf{V}_h$ and $p_h \in W_h$ such that

$$(2.3) \qquad \begin{array}{rcll} a(\underline{u}_h, \underline{\chi}) + b(\underline{\chi}, p_h) & = & 0 & \text{for all } \underline{\chi} \in \mathbf{V}_h, \\ b(\underline{u}_h, w) & = & -(f, w) & \text{for all } w \in W_h. \end{array}$$

This is a discrete linear system, and condition (2.2) implies, in particular, that the solution is unique.

We introduce the following operators $A = A_h : \mathbf{V}_h \to \mathbf{V}_h$ defined by

$$(A_h \underline{\chi}, \underline{\varphi}) = a(\underline{\chi}, \underline{\varphi}) \quad \text{for all } \underline{\chi} \text{ and } \underline{\varphi} \in \mathbf{V}_h;$$

$B = B_h : \mathbf{V}_h \to W_h$ defined by

$$(2.4) \qquad (B_h \underline{\chi}, w) = b(\underline{\chi}, w) \quad \text{for all } \underline{\chi} \in \mathbf{V}_h, \text{ and all } w \in W_h;$$

and $B^* = B_h^* : W_h \to \mathbf{V}_h$ defined by

$$(B_h^* w, \underline{\chi}) = b(\underline{\chi}, w) \quad \text{for all } w \in W_h \text{ and all } \underline{\chi} \in \mathbf{V}_h.$$

We observe that A is symmetric and positive definite and that B^* is the adjoint of B with respect to the L^2-inner products on $\mathbf{V}_h$ and W_h. Furthermore, the operators B and B^* are discrete analogs of the negative divergence $(-\nabla\cdot)$ and the gradient (∇) operators, respectively.

The discrete problem (2.3) is a linear system of the following form:

$$(2.5) \qquad \begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u}_h \\ p_h \end{pmatrix} = \begin{pmatrix} 0 \\ -Qf \end{pmatrix}.$$

Here, $Q = Q_h$ is the L^2- projection of $L^2(\Omega)$ onto the finite element space W_h. Note that there is no continuity requirement for the finite element functions in W_h across the element boundaries of $\mathcal{T}_h$. In practice, one chooses the space W_h consisting of discontinuous piecewise polynomials. In such cases it is clear that the projection Q_h is a local projection.

We observe that if $L = L_h : W_h \to W_h$ denotes the operator $L = BA^{-1}B^*$, then the solution p_h of (2.5) solves the reduced system

$$Lp_h = Qf,$$

which has the unique solution $p_h = L^{-1}Qf$.

To be more specific, consider the following pair of spaces due to Raviart and Thomas [30]. We denote by $Q_{r,s}(D)$, $r, s \geq 0$, the set of polynomials in two variables of degree r in the first variable and of degree s in the second variable, restricted to the given domain D. Also, $P_r(D)$ denotes the set of polynomials of two variables of total degree r restricted to D. Finally, let $P_r(E)$, for a one–dimensional boundary E, be the set of polynomials of one variable of degree r restricted to E. If $r \geq 0$ is fixed and $\mathcal{T}_h$ defines a rectangular grid, the space $\mathbf{V}_h$ consists of piecewise polynomials which for each $T \in \mathcal{T}_h$ are in $Q_{r+1,r}(T) \times Q_{r,r+1}(T)$, while W_h consists of discontinuous functions in $Q_{r,r}(T)$. On the other hand, if $\mathcal{T}_h$ consists of triangles, $\mathbf{V}_h$ consists of piecewise elements of $P_r(T)^2 \cup \left\{ \begin{bmatrix} x_1\hat{v}_0 \\ x_2\hat{v}_0 \end{bmatrix} \right\}$, where $\hat{v}_0$ runs over the homogeneous polynomials of degree r, while the elements of W_h are discontinuous functions which are locally in $P_r(T)$. We observe that the requirement $\mathbf{V}_h \subset H(\text{div}; \Omega)$ implies that the normal component $\underline{\chi} \cdot \underline{n}$ of each $\underline{\chi} \in \mathbf{V}_h$ should be continuous across every edge of $\mathcal{T}_h$. Here and elsewhere $\underline{n}$ is a given unit vector normal to the given edge of the element. It is well known (cf. [30]) that the pairs of spaces $\mathbf{V}_h$ and W_h constructed above satisfy the Babuška–Brezzi condition (2.2). However, for the analysis below, more properties of these spaces are needed.

We stress that the analysis to follow is not restricted to the spaces defined above. Also, with minor modifications it should apply to domains of higher than two dimensions. What is important for the analysis is the validity of the assumptions stated below.

Assumption (A-I). Each element $\underline{\chi} \in \mathbf{V}_h$ can be defined locally on the basis of the following degrees of freedom:

$$(2.6)$$
$$\int_E \underline{\chi} \cdot \underline{n}w \, d\rho = F_E(w), \quad \text{for all } w|_E, \ w \in W(T) \text{ and all edges } E \text{ of } T,$$
$$\int_T \underline{\chi} \cdot \nabla w \, dx = G_T(\nabla w), \quad \text{for all } w \in W(T).$$

Here, $W(T)$ and $\mathbf{V}(T)$ denote the restrictions of the spaces W_h and $\mathbf{V}_h$ to any element $T \in \mathcal{T}_h$, $F_E : L^2(E) \to \mathbb{R}$ and $G_T : L^2(T) \to \mathbb{R}$ are given linear functionals, while ρ denotes the arc length along E. $\qquad \square$

The assumption (A-I) will be needed in order to establish the spectral equivalence given in Theorem 3.1 below.

This assumption implies that there is a positive constant c, independent of h, such that for any $T \in \mathcal{T}_h$

$$(2.7) \qquad \int_T |\underline{\chi}|^2 \, dx \leq c \left(h \sum_E \|F_E\|^2 + \|G_T\|^2 \right),$$

where the sum is taken over all edges E of T. Furthermore, $\| \cdot \|$ denotes the norm of F_E and G_T as linear functionals from $L^2(E)$ and $L^2(T)$, respectively, to $\mathbb{R}$. Note that since $w|_E$ can be chosen from two different sides of each edge E, the stability of the above procedure implies that the value of $\underline{\chi} \cdot \underline{n}|_E$ is specified solely on the functionals F_E, i.e., independently of G_T. Also, from the continuity property of $\underline{\chi} \cdot \underline{n}$ across E it follows that if

$$\int_E \underline{\chi} \cdot \underline{n} w \, d\rho = 0 \quad \text{for all } w \in W(T),$$

then $\underline{\chi} \cdot \underline{n}|_E \equiv 0$.

Another consequence of the solvability of (2.6) is that the degrees of freedom of the space $\{\nabla w : w \in W(T)\}$ are no more than the degrees of freedom of the space $\mathbf{V}_0(T) \equiv \{\underline{\chi} \in \mathbf{V}(T) : \underline{\chi} \cdot \underline{n} = 0 \text{ on the edges } E \text{ of } T\}$. Furthermore, there is a constant c, independent of h and w, such that

$$(2.8) \qquad \left(\int_T |\nabla w|^2 \, dx \right)^{\frac{1}{2}} \leq c \sup_{\underline{\chi} \in \mathbf{V}_0(T)} \frac{\int_T \underline{\chi} \cdot \nabla w \, dx}{\|\underline{\chi}\|_{0,T}}$$

for all $w \in W(T)$.

Assumption (A-II). For any $\underline{\chi} \in \mathbf{V}_h$ its divergence, $\nabla \cdot \underline{\chi}$ is in W_h. $\qquad \square$

The assumptions (A-I) and (A-II) imply the Babuška–Brezzi condition (2.2). We refer to [15] for this well-known result.

We remark that Theorem 3.1 below, which is proved by only using the main assumption (A-I), implies a version of the Babuška–Brezzi condition in proper mesh-dependent norms.

In order to carry out the construction of the additive Schwarz preconditioner in §4 below, two extra technical assumptions will be needed. We will assume that $\hat{\Omega}$ is a fixed *convex*, bounded, polygonal domain which is an extension of Ω. We emphasize that the domain $\hat{\Omega}$ is only needed in theoretical arguments and does not contribute to the construction of the computational method. Furthermore, if Ω itself is convex we can take $\hat{\Omega} = \Omega$ and the two assumptions below will be simplified.

An important tool in the construction of a Schwarz preconditioner is a proper coarse grid.

Assumption (A-III). There exists a coarse family of quasi-uniform triangulations $\mathcal{T}_H$ of Ω, with characteristic diameter $H > h$, such that $\mathcal{T}_h$ corresponds to a refinement of $\mathcal{T}_H$. Furthermore, $\mathcal{T}_h$ and $\mathcal{T}_H$ can be extended to quasiuniform triangulations $\hat{\mathcal{T}}_h$ and $\hat{\mathcal{T}}_H$ on $\hat{\Omega}$ with a corresponding refinement property. $\qquad \square$

Note that as a consequence of these extension properties we can define piecewise polynomial spaces $\hat{\mathbf{V}}_h$ and $\hat{W}_h$ on $\hat{\Omega}$ such that $\mathbf{V}_h$ and W_h correspond to restrictions of the extended spaces to the domain Ω.

We shall also assume $\hat{\mathbf{V}}_h$ and $\hat{W}_h$ are contained in higher-order piecewise polynomial spaces $\hat{\mathbf{V}}_h^1$ and $\hat{W}_h^1$, which is a stable pair with respect to the Babuška–Brezzi

condition (2.2), and such that the corresponding mixed method admits error estimates of order h^2 for solutions which are H^2-regular on $\hat{\Omega}$. With respect to these higher-order spaces on $\hat{\Omega}$ we define an operator $\hat{B} = \hat{B}_h : \hat{\mathbf{V}}_h^1 \to \hat{W}_h^1$ similar to the definition of B above (cf. (2.4)). Furthermore, $\hat{B}^*$ denotes the L^2-adjoint of $\hat{B}$. We also let $\hat{L} = \hat{B}\hat{B}^*$, and $\hat{Q} = \hat{Q}_h$ is the L^2-projection of $L^2(\hat{\Omega})$ onto $\hat{W}_h^1$.

With this notation introduced the last assumption can be formulated precisely.

Assumption (A-IV). The spaces $\hat{\mathbf{V}}_h$ and $\hat{W}_h$ are contained in spaces of higher-degree piecewise polynomials $\hat{\mathbf{V}}_h^1$ and $\hat{W}_h^1$ which satisfy the assumptions (A-I) and (A-II). Furthermore, if $\hat{p} \in H^2(\hat{\Omega}) \cap H_0^1(\hat{\Omega})$ and $\hat{p}_h = \hat{L}^{-1}\hat{Q}(-\Delta)\hat{p}$, then

$$\|\hat{p} - \hat{p}_h\|_{0,\hat{\Omega}} \le ch^2\|\hat{p}\|_{2,\hat{\Omega}}.$$

Here Δ denotes the Laplace operator and the constant c is independent of h and $\hat{p}$. $\square$

Of the four assumptions given above, (A-III) can be satisfied by a proper construction of the grid.

For the particular Raviart–Thomas spaces introduced above, depending on a degree parameter $r \ge 0$, the three other assumptions also hold. It is well known, and easy to verify, that (A-II) is satisfied. Assumption (A-IV) holds with $\hat{\mathbf{V}}_h^1 = \hat{\mathbf{V}}_h$ and $\hat{W}_h^1 = \hat{W}_h$ if $r \ge 1$ (cf. Falk and Osborn [23] or Brezzi and Fortin [15]). Then, since the lowest-order Raviart–Thomas spaces ($r = 0$) are contained in any higher-order space (e.g. $r = 1$), the assumption also holds when $r = 0$.

The verification of (A-I) is essentially done already by Raviart and Thomas [30]. For rectangular elements (2.6) takes the form

$$\int_E \underline{\chi} \cdot \underline{n}q \, d\rho = F_E(q), \quad \text{for all } q \in P_r(E) \text{ and all edges } E \text{ of } T,$$
$$\int_T \underline{\chi} \cdot \underline{q} \, dx = G_T(\underline{q}), \quad \text{for all } \underline{q} \in Q_{r-1,r}(T) \times Q_{r,r-1}(T).$$

These equations specify $4(r+1) + 2r(r+1) = 2(r+1)(r+2)$ degrees of freedom of $\underline{\chi} \in \mathbf{V}(T)$, i.e., precisely the degrees of freedom in $Q_{r+1,r}(T) \times Q_{r,r+1}(T)$. Similarly, for triangular Raviart–Thomas elements, (2.6) reads

$$\int_E \underline{\chi} \cdot \underline{n}q \, d\rho = F_E(q), \quad \text{for all } q \in P_r(E) \text{ and all edges } E \text{ of } T,$$
$$\int_T \underline{\chi} \cdot \underline{q} \, dx = G_T(\underline{q}), \quad \text{for all } \underline{q} \in P_{r-1}(T)^2.$$

These equations specify $3(r+1) + r(r+1) = (r+1)(r+3)$ degrees of freedom for $\underline{\chi} \in \mathbf{V}(T)$, which are exactly the degrees of freedom in $P_r(T)^2 \cup \left\{ \left[\begin{smallmatrix} x_1 \hat{v}_0 \\ x_2 \hat{v}_0 \end{smallmatrix} \right] \right\}$ where $\hat{v}_0$ runs over the homogeneous polynomials of degree r.

3. THE SPECTRALLY EQUIVALENT LOCAL OPERATOR

In the rest of this paper we will usually drop the subscript h which indicates that for example operators or spaces depend on the fine grid $\mathcal{T}_h$, while dependence on the coarse grid $\mathcal{T}_H$ will be explicitly indicated. The analysis done in this section only relies on the assumption (A-I).

Many different approaches have been suggested in order to construct iterative methods for systems of the form (2.5). In our computations, presented below in §5, we have used the minimum residual method, or the SYMLQ method, originally proposed by Paige and Saunders [28] for general symmetric, indefinite problems.

Alternative iterative methods for saddle point systems can e.g. be found in Bramble and Pasciak [7], Bramble, Pasciak and Vassilev [11], Verfürth [37], Elman and Golub [21], Queck [29], Bank, Welfert and Yserentant [5], Axelsson [2], Axelsson and Vassilevski [3, 4].

The minimum residual method generates approximations of the solution of the linear system in the Krylov space generated by the symmetric operator

$$\begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix}$$

defined on the product space $\mathbf{V} \times W$. Therefore, only actions of this operator have to be computed during the iterations, and, owing to the symmetry of the operator, short recurrence relations are obtained. The convergence rate of the minimum residual method applied to systems of the form (2.5) will depend on spectral properties of the positive definite operators A and $L = BA^{-1}B^*$. If each of these operators are well conditioned, and they are properly scaled relative to each other, fast convergence of the iterative procedure can be guaranteed. For discussions on the application of the minimum residual method to systems of the form (2.5) we refer to Rusten and Winther [31, 32], Silvester and Wathen [33] and Vassilevski and Lazarov [34].

By the assumptions on the coefficient matrix k it follows that there are positive constants $\alpha_0, \alpha_1 > 0$ such that

$$\alpha_0 I \le k^{-1}(x) \le \alpha_1 I \quad \text{for all} \ \ x \in \Omega,$$

where I denotes the 2×2 identity matrix. Hence, the operator A is well conditioned, independent of h, in the sense that

(3.1) $$\alpha_0(\underline{\chi}, \underline{\chi}) \le (A\underline{\chi}, \underline{\chi}) \le \alpha_1(\underline{\chi}, \underline{\chi}) \quad \text{for all} \ \ \underline{\chi} \in \mathbf{V}.$$

Therefore, the performance of the minimum residual method will essentially be governed by the operator $L = BA^{-1}B^*$, which we recall is a discrete, nonlocal analog of the differential operator $-\nabla \cdot (k\nabla)$. Furthermore, since A is spectrally equivalent to the identity operator on $\mathbf{V}$, L is spectrally equivalent to BB^*, which is a nonlocal discrete analog of the negative Laplace operator.

The stability condition (2.2) will imply that

$$(BB^*w, w) = \sup_{\underline{\chi} \in \mathbf{V}} \frac{(B^*w, \underline{\chi})^2}{\|\underline{\chi}\|_0^2} \ge \sup_{\underline{\chi} \in \mathbf{V}} \frac{b(\underline{\chi}, w)^2}{\|\underline{\chi}\|_{H(\text{div};\Omega)}^2} \ge \beta_0^2 \|w\|_0^2.$$

Hence, the smallest eigenvalue of BB^* is bounded below, independent of h. However, since BB^* is a discrete analog of the negative Laplace operator, the largest eigenvalue will grow proportionally to h^{-2}. Consequently, a preconditioner for BB^* is necessary in order to obtain an efficient iterative method.

Following the ideas from [32], we introduce a positive definite operator N on W and consider the preconditioned system

(3.2) $$\begin{pmatrix} A & B^* \\ N^{-1}B & 0 \end{pmatrix} \begin{pmatrix} \underline{u}_h \\ p_h \end{pmatrix} = \begin{pmatrix} 0 \\ -N^{-1}Qf \end{pmatrix},$$

which is equivalent to (2.5). Furthermore, the coefficient operator is symmetric if the L^2-inner product on W is replaced by $(N\cdot, \cdot)$, and, if the preconditioner N is chosen such that the condition number of $N^{-1}BB^*$ is independent of h, the minimum residual method applied to (3.2) converges with a rate independent of h.

We remark that in order to compute the action of the coefficient matrix of (3.2) we need to compute the action of N^{-1}, while the action of N is never required.

We should emphasize here that every effective iterative method for the system (2.5) requires a preconditioner for the operator L, or BB^*. Hence, our discussion of preconditioners is not limited to the minimum residual method. For a discussion of the relations between several of the suggested iterative methods for saddle point problems we refer to Bramble and Pasciak [8].

The main purpose of the present paper is to construct a family of domain decomposition preconditioners for the discrete Laplace operator BB^*, i.e., preconditioners constructed by solving corresponding problems on overlapping subdomains. However, if we apply this approach directly to the nonlocal operator BB^*, the subdomain problems will be nonlocal, which obviously is not desirable. However, we will show below that BB^* is spectrally equivalent to a local operator. This local operator will then be used to define the preconditioners.

Let $\mathcal{E}$ denote the set of all edges of elements in $\mathcal{T}$. For each edge $E \in \mathcal{E}$ let $n = n_E(x)$ be a fixed unit vector normal to E, and for $w \in W$ let $[w](x) = [w]_E(x)$ denote the jump of w, i.e.

$$[w](x) = w(x + 0n) - w(x - 0n).$$

Note, that by extending w by zero outside Ω, $[w]$ is also defined on boundary edges. We define the bilinear form $\mathcal{A} = \mathcal{A}_h$ on $W \times W$ by

$$\mathcal{A}(w, \phi) = \sum_{T \in \mathcal{T}} \int_T \nabla w \cdot \nabla \phi \, dx + \sum_{E \in \mathcal{E}} h^{-1} \int_E [w][\phi] \, d\rho.$$

We observe that this bilinear form corresponds to the interior penalty discretization of the negative Laplace operator with a homogeneous Dirichlet boundary condition (cf. [1]). Note also that since the boundary of Ω is included in $\mathcal{E}$, the form $\mathcal{A}$ is positive definite on W. Furthermore, the bilinear form $\mathcal{A}$, and hence the associated operator, is local with respect to w and ϕ. For convenience, we will also let $\mathcal{A}$ denote this associated operator on W, i.e., $\mathcal{A} : W \mapsto W$ is the positive definite operator

$$(\mathcal{A}w, \phi) = \mathcal{A}(w, \phi) \quad \text{for all } w, \phi \in W.$$

The following equivalence result will be established.

Theorem 3.1. *There exists a positive constant γ_0, independent of h, such that*

$$(3.3) \qquad \gamma_0^{-1} \mathcal{A}(w, w) \leq (B^*w, B^*w) \leq \gamma_0 \mathcal{A}(w, w) \quad \text{for all } w \in W.$$

Proof. Let $w \in W$ be given. We first establish the right inequality. To do this, we let $\underline{\chi} = -B^*w$, use the definitions of B^* and $b(\cdot, \cdot)$, and an integration by parts to obtain

$$(B^*w, B^*w) = \sum_{T \in \mathcal{T}} \int_T w \nabla \cdot \underline{\chi} \, dx = -\sum_{T \in \mathcal{T}} \int_T \underline{\chi} \cdot \nabla w \, dx + \sum_{E \in \mathcal{E}} \int_E \underline{\chi} \cdot n[w] \, d\rho.$$

Repeated use of the Cauchy-Schwarz inequality implies

$$\sum_{T \in \mathcal{T}} -\int_T \underline{\chi} \cdot \nabla w \, dx \leq \sum_{T \in \mathcal{T}} \|\underline{\chi}\|_{0,T} \|\nabla w\|_{0,T} \leq \|\underline{\chi}\|_0 \left(\sum_{T \in \mathcal{T}} \|\nabla w\|_{0,T}^2 \right)^{1/2}$$

and

$$\sum_{E \in \mathcal{E}} \int_E \underline{\chi} \cdot n[w] \, d\rho \leq \sum_{E \in \mathcal{E}} \|\underline{\chi} \cdot \underline{n}\|_{0,E} \|[w]\|_{0,E}$$

$$\leq \left(\sum_{E \in \mathcal{E}} h \|\underline{\chi} \cdot \underline{n}\|_{0,E}^2 \right)^{1/2} \left(\sum_{E \in \mathcal{E}} h^{-1} \|[w]\|_{0,E}^2 \right)^{1/2}.$$

Furthermore, the inverse inequality $h\|\underline{\chi} \cdot \underline{n}\|_{0,E}^2 \leq c\|\underline{\chi}\|_{0,T}^2$, where E is an edge of T, gives

$$\sum_{E \in \mathcal{E}} h \|\underline{\chi} \cdot \underline{n}\|_{0,E}^2 \leq c\|\underline{\chi}\|_0^2.$$

Hence, since $\underline{\chi} = -B^* w$ we have shown that there is a constant c, independent of h, such that

$$\|B^* w\|_0^2 \leq c\|B^* w\|_0 \mathcal{A}(w,w)^{1/2},$$

and this implies the right inequality of (3.3).

Next, we establish the left inequality of (3.3). First we bound the sum over the elements, then the sum over the edges. Let $T \in \mathcal{T}$ be given. Note that by extending elements of $\mathbf{V}_0(T) = \{\underline{\chi} \in \mathbf{V}(T) : \underline{\chi} \cdot \underline{n} = 0 \text{ on } \partial T\}$ by zero outside T, this space can be identified with a subspace of $\mathbf{V}$. Using (2.8) and integration by parts, we obtain

$$\|\nabla w\|_{0,T} \leq c \sup_{\underline{\chi} \in \mathbf{V}_0(T)} \frac{b(\underline{\chi}, w)}{\|\underline{\chi}\|_{0,T}} = c \sup_{\underline{\chi} \in \mathbf{V}_0(T)} \frac{(\underline{\chi}, B^* w)}{\|\underline{\chi}\|_{0,T}} \leq c \, \|B^* w\|_{0,T}$$

and the desired result follows by summing over the elements.

Finally, we bound the sum over the edges. By (2.6), for any given $w \in W$, we can define $\underline{\chi} \in \mathbf{V}$, element by element, such that

$$\sum_{E \in \mathcal{E}} \int_E \underline{\chi} \cdot \underline{n}[v] \, d\rho - \sum_{T \in \mathcal{T}} \int_T \underline{\chi} \cdot \nabla v \, dx = \sum_{E \in \mathcal{E}} \int_E [w][v] \, d\rho \quad \text{for all } v \in W,$$

and (2.7) implies that

$$(3.4) \qquad \|\underline{\chi}\|_0^2 \leq c \sum_{E \in \mathcal{E}} h \int_E [w]^2 \, d\rho.$$

Using the definition of B^*, integration by parts, and the properties of $\underline{\chi}$, we obtain

$$-(B^* w, \underline{\chi}) = -b(\underline{\chi}, w) = \sum_{E \in \mathcal{E}} \int_E [w]\underline{\chi} \cdot \underline{n} \, d\rho - \sum_{T \in \mathcal{T}} \int_T \underline{\chi} \cdot \nabla v \, dx = \sum_{E \in \mathcal{E}} \int_E [w]^2 \, d\rho.$$

However, together with (3.4), this implies that

$$\sum_{E \in \mathcal{E}} h^{-1} \int_E [w]^2 \, d\rho \leq c\|B^* w\|_0 \left(\sum_{E \in \mathcal{E}} h^{-1} \int_E [w]^2 \, d\rho \right)^{\frac{1}{2}},$$

and the desired result follows.   □

The mesh-dependent norm associated with the form $\mathcal{A}$ will be denoted $|\cdot|_{1,h}$, i.e.,

$$|w|_{1,h} = \sqrt{\mathcal{A}(w,w)}.$$

Observe that Theorem 3.1 implies a Babuška–Brezzi condition of the form

$$\gamma_0^{-1}|w|_{1,h} \leq \sup_{\underline{\chi} \in \mathbf{V}} \frac{b(\underline{\chi}, w)}{\|\underline{\chi}\|_0} \quad \text{for all } w \in W.$$

Observe also that the proof of the theorem above only depends on assumption (A-I). Hence, if (A-II)–(A-IV) hold, the proof also applies to the operator $\hat{B}^*$ defined from the spaces $\hat{\mathbf{V}}^1$ and $\hat{W}^1$. In particular, the following corollary will be useful below.

Corollary 3.1. *Assume that the additional assumptions* (A-II)–(A-IV) *are satisfied. Let* $v \in W$, *and let* $\hat{v} \in \hat{W} \subset \hat{W}^1$ *be the extension of* v *by zero outside* Ω. *Then the two norms*

$$|v|_{1,h} \quad and \quad \|\hat{B}^*\hat{v}\|_{0,\hat{\Omega}}$$

are equivalent, independently of h.

4. THE ADDITIVE SCHWARZ METHOD

In the rest of this paper we assume that all the four assumptions (A-I)–(A-IV) hold.

It follows from Theorem 3.1 above that the interior penalty method can be used to construct effective preconditioners for the mixed system (2.5). As an application of this result we shall in this section define and analyze an additive Schwarz preconditioner based on the local form $\mathcal{A}$. The analysis relies on the construction of a specific coarse space W_H. We remark that part of the analysis below is rather similar to what is done in the analysis of Schwarz preconditioners for conforming finite element methods. However, the analysis related to the coarse space, cf. Lemma 4.2 below, has additional difficulties caused by the nonconformity of the spaces.

Let $\{\Omega_j\}_{j=1}^J$ be the elements of the coarse triangulation $\mathcal{T}_H$. By extending each Ω_j to a larger domain Ω_j', we obtain an overlapping covering of Ω. The domains Ω_j' are also assumed to be mesh domains, i.e., the boundaries consist of element edges. We assume throughout the paper that there is a constant β_1, independent of H, such that

$$(4.1) \qquad \text{dist}(\partial\Omega_j', \partial\Omega_j) \geq \beta_1 H \quad \text{for } j = 1, 2, \dots, J.$$

This condition expresses that the overlap is sufficiently large. On the other hand, we also need to limit the maximum number of overlaps. Let κ_j denote the characteristic functions on $\overline{\Omega}_j''$, where Ω_j'' is the union of Ω_j' and all its neighbor elements in $\mathcal{T}_h$, and let

$$\kappa = \sum_{j=1}^J \kappa_j.$$

We then assume that

$$(4.2) \qquad \sup_{x \in \overline{\Omega}} \kappa(x) \leq \beta_2,$$

where β_2 is independent of H. In practice, β_2 can usually be taken to be 4.

We define subspaces W_j of W by

$$W_j = \{v \in W : \text{supp } v \subset \overline{\Omega}_j'\} \quad \text{for } j = 1, 2, \dots, J.$$

Furthermore, the projection of W onto W_j with respect to the form $\mathcal{A}$ is denoted by P_j, i.e., $P_j : W \mapsto W_j$ is determined by

$$\mathcal{A}(P_j v, w) = \mathcal{A}(v, w) \quad \text{for all } w \in W_j.$$

As is the case for conforming finite element methods, the construction of a Schwarz preconditioner also requires a proper coarse space. Recall that W^c denotes the subspace of W consisting of piecewise constants with respect to $\mathcal{T}$. Also, let $S_H \subset H_0^1(\Omega)$ be the conforming finite element space associated with the triangulation $\mathcal{T}_H$ of piecewise linear (or bilinear) polynomials. The coarse space W_H is defined by

$$W_H = \{Q^c \psi \ : \psi \in S_H\},$$

where Q^c is the L^2–projection onto the space W^c. Furthermore, $P_H : W \mapsto W_H$ is the projection onto W_H with respect to the form $\mathcal{A}$. Frequently below we will write P_0 instead of P_H.

Let $P = P_H + \sum_{j=1}^{J} P_j = \sum_{j=0}^{J} P_j$ and define the preconditioner $N^{-1} : W \mapsto W$ by

$$N^{-1} = P\mathcal{A}^{-1}.$$

Then the action of N^{-1} can be computed by solving a global problem on W_H and a sequence of possible parallel local problems. Furthermore, the operator $N^{-1}\mathcal{A}$, which essentially determines the convergence properties of the minimum residual method, is equal to P.

The following result shows that the condition number of P is bounded independently of h and H.

Theorem 4.1. *There exists a positive constant γ_1, independent of h and H, such that*

(4.3) $$\gamma_1^{-1} \mathcal{A}(v, v) \leq \mathcal{A}(Pv, v) \leq \gamma_1 \mathcal{A}(v, v) \quad \text{for all } v \in W.$$

Proof. We first establish the upper bound of (4.3). First use the Cauchy–Schwarz inequality

(4.4) $$\mathcal{A}(Pv, v) \leq |Pv|_{1,h} |v|_{1,h}.$$

Then, if we let $\kappa_0 \equiv 1$, we obtain by the Cauchy–Schwarz inequality again

$$\mathcal{A}(Pv, Pv) = \sum_{i,j=0}^{J} \sum_{T \in \mathcal{T}} \int_T \nabla P_i v \cdot \nabla P_j v \, dx + \sum_{i,j=0}^{J} \sum_{E \in \mathcal{E}} h^{-1} \int_E [P_i v][P_j v] \, d\rho$$

$$\leq \sum_{i,j=0}^{J} \sum_{T \in \mathcal{T}} \int_T \kappa_j^2 |\nabla P_i v|^2 \, dx + \sum_{i,j=0}^{J} \sum_{E \in \mathcal{E}} h^{-1} \int_E \kappa_j^2 [P_i v]^2 \, d\rho$$

$$\leq (1 + \beta_2)^2 \sum_{i=0}^{J} \mathcal{A}(P_i v, P_i v) = (1 + \beta_2)^2 \mathcal{A}(Pv, v).$$

Together with inequality (4.4), this implies the desired uniform upper bound.

In order to show the lower bound of (4.3) we need two lemmas, which will be established below.

Lemma 4.1. *For any $v \in W$ there exists a decomposition $v = \sum_{j=1}^{J} v_j$, $v_j \in W_j$, satisfying*

(4.5) $$\sum_{j=1}^{J} \mathcal{A}(v_j, v_j) \leq c(H^{-2} \|v\|_0^2 + \mathcal{A}(v, v))$$

for some positive constant c independent of h and H.

Lemma 4.2. *There exists a proper decomposition of any $v \in W$ of the form $v = (v - v_H) + v_H$ where $v_H \in W_H$ satisfies the properties*

$$(4.6) \qquad \|v - v_H\|_0 \leq cH|v|_{1,h}$$

and

$$(4.7) \qquad |v_H|_{1,h} \leq c|v|_{1,h},$$

where the constant c is independent of h and H.

From these two lemmas the desired lower bound follows by standard arguments (cf., e.g., Lions [26], Matsokin and Nepomnyaschikh [27] or Dryja and Widlund [19]). By applying Lemma 4.2 to an arbitrary element $v \in W$ and thereafter using the decomposition of Lemma 4.1 on the element $v - v_H$, we obtain

$$v = v_H + \sum_{j=1}^{J} v_j,$$

where $v_H \in W_H, v_j \in W_j$ and

$$\mathcal{A}(v_H, v_H) + \sum_{j=1}^{J} \mathcal{A}(v_j, v_j) \leq \mathcal{A}(v_H, v_H) + c(H^{-2}\|v - v_H\|_0^2 + \mathcal{A}(v - v_H, v - v_H))$$

$$\leq c\mathcal{A}(v, v).$$

From the Cauchy–Schwarz inequality we then derive

$$\mathcal{A}(v, v) = \mathcal{A}(v, v_H) + \sum_{j=1}^{J} \mathcal{A}(v, v_j)$$

$$= \mathcal{A}(P_H v, v_H) + \sum_{j=1}^{J} \mathcal{A}(P_j v, v_j)$$

$$\leq \left(\mathcal{A}(v_H, v_H) + \sum_{j=1}^{J} \mathcal{A}(v_j, v_j) \right)^{\frac{1}{2}} \left(\sum_{j=0}^{J} \mathcal{A}(P_j v, v) \right)^{\frac{1}{2}}$$

$$\leq c(\mathcal{A}(v, v))^{\frac{1}{2}} (\mathcal{A}(Pv, v))^{\frac{1}{2}}.$$

This shows the lower bound of (4.3). In order to complete the proof of Theorem 4.1 we have to establish the Lemmas 4.1 and 4.2. This will be done below. $\square$

Proof of Lemma 4.1. Let $v \in W$ be given. The desired decomposition of v is constructed by a standard argument which is similar to what is used for conforming methods (cf. Dryja and Widlund [19] or Dryja, Smith and Widlund [20]). Because of the properties of the covering $\{\Omega_j'\}_{j=1}^{J}$ of Ω given above, there exists a partition of unity of smooth functions on Ω, $\{\theta_j\}_{j=1}^{J}$, such that

$$\sum_{j=1}^{J} \theta_j = 1, \quad 0 \leq \theta_j \leq 1 \text{ in } \Omega;$$

$$\text{supp } \theta_j \subset \overline{\Omega}_j';$$

$$(4.8) \qquad \|\nabla \theta_j\|_{L^\infty(\Omega)} \leq cH^{-1},$$

where c is independent of H.

Let I_h be a local interpolation operator defined for piecewise continuous functions (with respect to the elements of $\mathcal{T}$) into the space W. Such an operator will have the restriction property

$$(4.9) \qquad I_h(v)|_T = v|_T \quad \text{if} \quad v \in W(T),$$

and satisfy

$$(4.10) \qquad \|I_h(f)\|_{L^\infty(T)} \le c \, \|f\|_{L^\infty(T)},$$

where the constant c is independent of h. We now define $v_j \in W_j$ by $v_j = I_h(\theta_j v)$. From (4.9) it follows that $v = \sum_{j=1}^J v_j$. Hence, it remains to establish the inequality (4.5).

Define for convenience the sets of triangles $\mathcal{T}_j = \{T \in \mathcal{T} \, : \, T \subset \Omega'_j\}$ and the sets of edges $\mathcal{E}_j = \{E \in \mathcal{E} \, : \, E \subset \overline{\Omega}'_j\}$. Since $v_j \in W_j$, we obviously have

$$\mathcal{A}(v_j, v_j) = \sum_{T \in \mathcal{T}_j} \int_T |\nabla v_j|^2 \, dx + \sum_{E \in \mathcal{E}_j} h^{-1} \int_E [v_j]^2 \, d\rho.$$

We estimate now each term in this expression separately. Let $T \in \mathcal{T}_j$ be fixed, and let $\hat{\theta}_j$ be the average value of θ_j on T. We then have

$$\int_T |\nabla I_h(\theta_j v)|^2 \, dx \le 2 \int_T |\nabla I_h((\theta_j - \hat{\theta}_j)v)|^2 \, dx + 2 \int_T \hat{\theta}_j^2 |\nabla v|^2 \, dx,$$

where we have used (4.9). From (4.10) we derive

$$\int_T |\nabla I_h((\theta_j - \hat{\theta}_j)v)|^2 \, dx \le c\|I_h((\theta_j - \hat{\theta}_j)v)\|_{L^\infty(T)}^2 \le c\|(\theta_j - \hat{\theta}_j)v\|_{L^\infty(T)}^2.$$

Since, by the approximation property of the mean value and (4.8),

$$\|\theta_j - \hat{\theta}_j\|_{L^\infty(T)} \le ch\|\nabla \theta_j\|_{L^\infty(T)} \le chH^{-1}$$

and

$$\|v\|_{L^\infty(T)} \le ch^{-1}\|v\|_{0,T},$$

we obtain the local estimate

$$(4.11) \qquad \int_T |\nabla v_j|^2 \, dx \le c(H^{-2}\|v\|_{0,T}^2 + \int_T |\nabla v|^2 \, dx).$$

The edge integrals are estimated similarly. Let $E \in \mathcal{E}_j$ and let T^+ and T^- be the two neighboring elements in $\mathcal{T}$. By arguments similar to the ones given above we easily establish

$$(4.12) \qquad h^{-1} \int_E [v_j]^2 \, d\rho \le c(H^{-2}\|v\|_{0,T^+ \cup T^-}^2 + h^{-1} \int_E [v]^2 \, d\rho).$$

The desired estimate (4.5) follows by summing up the local estimates (4.11) and (4.12). For convenience we introduce the set $\mathcal{T}_j' = \{T \in \mathcal{T} : T \text{ is a neighboring ele-}$

ment to $E \in \mathcal{E}_j$}. We then have

$$\sum_{j=1}^{J} \mathcal{A}(v_j, v_j) \leq \sum_{j=1}^{J} \left(\sum_{T \in \mathcal{T}_j} \int_T |\nabla v_j|^2 \, dx + \sum_{E \in \mathcal{E}_j} h^{-1} \int_E [v_j]^2 \, d\rho \right)$$

$$\leq c \sum_{j=1}^{J} \left(\sum_{T \in \mathcal{T}_j'} (H^{-2}\|v\|_{0,T}^2 + \int_T |\nabla v|^2 \, dx) + \sum_{E \in \mathcal{E}_j} h^{-1} \int_E [v]^2 \, d\rho \right)$$

$$\leq c(H^{-2}\|v\|_0^2 + \mathcal{A}(v,v)).$$

Here the final constant c is independent of h and H, but it depends on the number of overlaps β_2. □

Proof of Lemma 4.2. Let $\hat{S}_H \subset H_0^1(\hat{\Omega})$ be the conforming finite element space associated with the triangulation $\hat{\mathcal{T}}_H$ of piecewise linear (or bilinear) polynomials. Hence, by extending elements of S_H by zero outside Ω, S_H can be identified with a subspace of $\hat{S}_H$. Let $\hat{Z}$ be the set of interior nodes corresponding to the triangulation $\hat{\mathcal{T}}_H$, and let $\{\psi_z\}_{z \in \hat{Z}}$ be the corresponding nodal basis of $\hat{S}_H$ such that any $\phi \in \hat{S}_H$ has the representation

$$\phi(x) = \sum_{z \in \hat{Z}} \phi(z)\psi_z(x).$$

If $Z \subset \hat{Z}$ denotes the set of interior nodal points in $\mathcal{T}_H$, we define a "restriction map" $R_H : \hat{S}_H \mapsto S_H$ by

$$(R_H \phi)(x) = \sum_{z \in Z} \phi(z)\psi_z(x).$$

By utilizing the equivalence between the L^2-norm and the discrete norm

$$\left(H^2 \sum_{z \in \hat{Z}} \phi^2(z) \right)^{1/2}$$

on $\hat{S}_H$, we obtain the estimate

(4.13) $$\|\phi - R_H\phi\|_{0,\hat{\Omega}} \leq c\|\phi\|_{0,\hat{\Omega}\backslash\Omega},$$

which will be useful below. By a similar equivalence between the $H^1(\hat{\Omega})$ norm and a corresponding discrete finite difference norm obtained from the nodal values we also obtain

(4.14) $$\|R_H\phi\|_1 \leq c(\|\phi\|_{1,\hat{\Omega}} + H^{-1}\|\phi\|_{0,\hat{\Omega}\backslash\Omega}).$$

We introduce the L^2-projection of $L^2(\hat{\Omega})$ into $\hat{S}_H$ which will be denoted by $\hat{Q}_H$.

Let $v \in W$ be given. In order to define the desired element $v_H \in W_H$, we first let $\hat{v} \in \hat{W} \subset \hat{W}^1$ be the extension of v by zero outside Ω. Recall that $\hat{L} = \hat{B}\hat{B}^* : \hat{W}^1 \mapsto \hat{W}^1$ is an approximation of the negative Laplace operator on $\hat{\Omega}$. Note, in particular, that it follows from Corollary 3.1 that

(4.15) $$\|\hat{B}^*\hat{v}\|_{0,\hat{\Omega}} \leq c|v|_{1,h}$$

and hence, by an inverse inequality,

(4.16) $$\|\hat{L}\hat{v}\|_{0,\hat{\Omega}} \leq ch^{-1}|v|_{1,h}.$$

Here the constant c is independent of h.

Since $\hat{\Omega}$ is convex and $\hat{L}\hat{v} \in L^2(\hat{\Omega})$, it follows from elliptic regularity that there is a unique $\hat{p} \in H^2(\hat{\Omega}) \cap H_0^1(\hat{\Omega})$ such that

$$(4.17) \qquad\qquad -\Delta \hat{p} = \hat{L}\hat{v} \quad \text{in} \quad \hat{\Omega}.$$

Furthermore, from (4.16) we derive that there is a constant c independent of h such that

$$(4.18) \qquad\qquad \|\hat{p}\|_{2,\hat{\Omega}} \leq ch^{-1}|v|_{1,h}.$$

By multiplying (4.17) by $\hat{p}$ and utilizing that $\hat{L}\hat{v} = \nabla \cdot (\hat{B}^*\hat{v})$, we also obtain from (4.15) that

$$(4.19) \qquad\qquad \|\nabla \hat{p}\|_{0,\hat{\Omega}} \leq c|v|_{1,h}.$$

Note that it follows from (4.17) that $\hat{v} = \hat{L}^{-1}\hat{Q}(-\Delta)\hat{p}$. Hence, by assumption **(A-IV)** and (4.18),

$$\|\hat{v} - \hat{p}\|_{0,\hat{\Omega}} \leq ch^2\|\hat{p}\|_{2,\hat{\Omega}} \leq ch|v|_{1,h}.$$

Then we also get the estimates in the subdomains Ω and $\hat{\Omega} \setminus \Omega$,

$$(4.20) \qquad\qquad \|v - \hat{p}\|_{0,\Omega} \leq ch|v|_{1,h},$$

and, since $\hat{v} = 0$ outside Ω,

$$\|\hat{p}\|_{0,\hat{\Omega}\setminus\Omega} \leq ch|v|_{1,h}.$$

Define now $v_H = Q^c R_H \hat{Q}_H \hat{p}$. In order to derive the proper estimates for $\|v - v_H\|_0$ and $|v_H|_{1,h}$, we first observe that it follows from (4.19) and the approximation property of $\hat{S}_H$ that

$$(4.21) \quad \|\hat{Q}_H \hat{p}\|_{0,\hat{\Omega}\setminus\Omega} \leq c(\|\hat{p}\|_{0,\hat{\Omega}\setminus\Omega} + \|(I - \hat{Q}_H)\hat{p}\|_{0,\hat{\Omega}\setminus\Omega})$$
$$\leq c(\|\hat{p}\|_{0,\hat{\Omega}\setminus\Omega} + H\|\nabla\hat{p}\|_{0,\hat{\Omega}}) \leq c(h + H)|v|_{1,h} \leq cH|v|_{1,h},$$

where c is independent of h and H. Hence, we derive from (4.13) and (4.14), and by using the approximation property of $\hat{S}_H$ and (4.19) once more, that

$$\|\hat{p} - v_H\|_0 \leq \|\hat{p} - \hat{Q}_H\hat{p}\|_{0,\hat{\Omega}} + \|(I - R_H)\hat{Q}_H\hat{p}\|_{0,\hat{\Omega}} + \|(I - Q^c)R_H\hat{Q}_H\hat{p}\|_0$$
$$\leq c(H\|\nabla\hat{p}\|_{0,\hat{\Omega}} + \|\hat{Q}_H\hat{p}\|_{0,\hat{\Omega}\setminus\Omega})$$
$$\leq cH|v|_{1,h}.$$

However, together with (4.20) this implies the estimate (4.6).

Finally, we have to show (4.7), i.e.,

$$|v_H|_{1,h} = |Q^c R_H \hat{Q}_H \hat{p}|_{1,h} \leq c|v|_{1,h}.$$

From (4.14), (4.19), (4.21) and the fact that $\hat{Q}_H$ is uniformly bounded in H^1 it follows that

$$\|R_H \hat{Q}_H \hat{p}\|_1 \leq c(\|\hat{Q}_H\hat{p}\|_{1,\hat{\Omega}} + H^{-1}\|\hat{Q}_H\hat{p}\|_{0,\hat{\Omega}\setminus\Omega}) \leq c|v|_{1,h}.$$

Therefore, it is enough to show that

$$(4.22) \qquad\qquad |Q^c\phi|_{1,h} \leq c\|\phi\|_1,$$

for any $\phi \in S_H$. However, from the quasi-uniformity of $\mathcal{T}_h$, and since $Q^c\phi$ is piecewise constant, we have

$$|Q^c\phi|_{1,h}^2 = \sum_{E\in\mathcal{E}} h^{-1} \int_E [Q^c\phi]^2 \, d\rho \le c \sum_{E\in\mathcal{E}} [Q^c\phi]^2.$$

Furthermore, for any $E \in \mathcal{E}$ and any $\phi \in S_H$ we obtain, by adding and subtracting values of ϕ on E,

$$[Q^c\phi]^2 \le ch^{-4}(\int_{T^+} \phi(x)\, dx - \int_{T^-} \phi(x)\, dx)^2 \le ch^{-4}(\int_{T^+\cup T^-} e(x) \cdot \nabla\phi(x)\, dx)^2,$$

where $|e(x)|$ is the distance from x to E and T^+ and T^- are the two neighboring elements sharing E as a common boundary. Since

$$\int_{T^+\cup T^-} |e(x)|^2 \, dx \le ch^4,$$

we obtain from the Cauchy–Schwarz inequality that

$$[Q^c\phi]^2 \le c \int_{T^+\cup T^-} |\nabla\phi|^2 \, dx,$$

and hence (4.22) is obtained by summing this inequality over all $E \in \mathcal{E}$. This completes the proof of Lemma 4.2.   □

5. Numerical results

   The purpose of this section is to present some results of numerical experiments which seem to conform the theoretical results derived in this paper.

   Throughout this section Ω will be taken to be the unit square, $(0,1) \times (0,1)$. Furthermore, the spaces $\mathbf{V}$ and W will be the rectangular Raviart-Thomas spaces of order zero. The triangulations $\mathcal{T}_h$ and $\mathcal{T}_H$ are uniform and consist of rectangles of size $h \times h$ and $H \times H$, respectively, while the overlapping domains Ω_j' are of size $2H$.

   In the examples below we will apply the preconditioned minimum residual method to the discrete system (2.5). The additive Schwarz preconditioner N is constructed as indicated in §4 above. In particular, we shall be interested in the number of iterations required to reach the tolerance

(5.1) $$\|r_k\|/\|r_0\| \le 10^{-4},$$

and the condition number of the operator $N^{-1}BB^*$. Here, r_k is the residual in iteration k and, with $r_k = (\chi, w)$,

$$\|r_k\| = (\|\chi\|_0^2 + (Nw, w))^{1/2}.$$

Example 5.1. In the first example we choose $k = I$, where I is the 2×2 identity matrix, and $f = 1$. In this case $\alpha_0 = \alpha_1 = 1$ in the inequalities (3.1). In Table 1 (next page) we list the number of iterations needed to reach the tolerance (5.1) for different choices of h and H. Also, an estimate of the condition number of the operator $N^{-1}BB^*$ is given in parentheses. The condition numbers are estimated using the conjugate gradient method. The table indicates clearly that the condition number of $N^{-1}BB^*$ is bounded independently of h and H. In fact, the eigenvalues of the operator $N^{-1}BB^*$ are contained in the interval $[0.95, 4]$, hence we expect that the number of iterations of the preconditioned minimum residual method is uniformly bounded. The iteration counts in Table 1 confirm that this is the case.

TABLE 1. Iteration counts and condition numbers for Example 5.1

	$h = 1/8$	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$
$H = 1/4$	14 (4.0)	16 (4.0)	18 (4.0)	16 (4.0)	16 (4.1)
$H = 1/8$		14 (3.7)	14 (4.0)	16 (4.0)	16 (4.0)
$H = 1/16$			12 (4.0)	14 (4.0)	14 (4.0)
$H = 1/32$				12 (3.7)	12 (4.0)

TABLE 2. Iteration counts for Example 5.2

	$h = 1/8$	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$
$H = 1/4$	70	84	90	92	92
$H = 1/8$		85	91	93	94
$H = 1/16$			90	96	97
$H = 1/32$				93	95

Example 5.2. In the second example we choose

$$k(x) = \begin{pmatrix} 1 + 4(x_1^2 + x_2^2) & 3x_1x_2 \\ 3x_1x_2 & 1 + 11(x_1^2 + x_2^2) \end{pmatrix}.$$

Here the inequalities (3.1) hold with $\alpha_0 = 1$ and $\alpha_1 = 25$. We use the same preconditioner as in Example 5.1. Hence, we do not take the variation of k into account in the construction of the preconditioner. We observe that the preconditioned minimum residual method requires more iterations to converge, cf. Table 2. However, the number of iterations still seems to be uniformly bounded. We remark that the actual number of iterations would be reduced if the variable coefficient k was included in the preconditioner.

REFERENCES

1. D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal. **19** (1982), pp. 742–760. MR **83f:**65173
2. O. Axelsson, *Numerical algorithms for indefinite problems*, Elliptic Problem Solvers II (G. Birkhoff and A. Schoenstadt, eds.), Academic Press, Orlando, 1984, pp. 219–232. CMP 17:03
3. O. Axelsson and P. S. Vassilevski, *A black box generalized conjugate gradient solver with inner iterations and variable–step preconditioning*, SIAM J. Matrix Anal. Appl. **12** (1991), pp. 625–644. MR **92m:**65046
4. _____, *Construction of variable–step preconditioners for inner–outer iterative methods*, Iterative Methods in Linear Algebra (R. Beauwens and P. de Groen, eds.), North-Holland, Amsterdam, 1992, pp. 1–14. CMP 92:11
5. R. E. Bank, B. D. Welfert, and H. Yserentant, *A class of iterative methods for solving saddle point problems*, Numer. Math. **56** (1990), pp. 645–666. MR **91b:**65035
6. P. E. Bjørstad and O. B. Widlund, *Iterative methods for the solution of elliptic problems on regions partitioned into substructures*, SIAM J. Numer. Anal. **23** (1986), pp. 1097–1120. MR **88h:**65188
7. J. H. Bramble and J. E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp. **50** (1988), pp. 1–17. MR **89m:**65097a
8. _____, *Iterative techniques for time dependent Stokes problem.* Preprint 1994.

9. J. H. Bramble, J. E. Pasciak, and A. H. Schatz, *The construction of preconditioners for elliptic problems by substructuring. I*, Math. Comp. **47** (1986), pp. 103–134. MR **87m:**65174

10. ———, *An iterative method for elliptic problems on regions partitioned into substructures*, Math. Comp. **46** (1986), pp. 361–369. MR **88a:**65123

11. J. H. Bramble, J. E. Pasciak, and A. T. Vassilev, *Analysis of the inexact Uzawa algorithm for saddle point problems*. Preprint 1994.

12. J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms*, Math. Comp. **56** (1991), pp. 1–34. MR **91h:**65159

13. F. Brezzi, J. Douglas, Jr., M. Fortin, and L. D. Marini, *Efficient rectangular mixed finite elements in two and three space variables*, R.A.I.R.O. Mathematical Modelling and Numerical Analysis **21** (1987), no. 4, pp. 581–604. MR **88j:**65249

14. F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47** (1985), pp. 217–235. MR **87g:**65133

15. F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991. MR **92d:**65187

16. L. C. Cowsar, *Dual variable Schwarz methods for mixed finite elements*, Report TR93-09, Rice University, Houston, 1993.

17. L. C. Cowsar, J. Mandel, and M. F. Wheeler, *Balancing domain decomposition for mixed finite elements*, Math. Comp. **64** (1995), pp. 989–1015. MR **95j:**65161

18. J. Douglas, Jr. and J. Wang, *A new family of mixed finite element spaces over rectangles*, Comp. Appl. Math. **12** (1993), pp. 183–197. CMP 94:16

19. M. Dryja and O. B. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*, Technical Report 339, Courant Institute of Mathematical Sciences, 1987.

20. M. Dryja, B. Smith, and O. B. Widlund, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, Technical Report 638, Courant Institute of Mathematical Sciences, 1993.

21. H. C. Elman and G. Golub, *Inexact and preconditioned Uzawa algorithms for saddle point problems*, SIAM J. Numer. Anal. **31** (1994), pp. 1645–1661. MR **95f:**65065

22. R. E. Ewing and M. F. Wheeler, *Computational aspects of mixed finite element methods*, Numerical Methods for Scientific Computing (R. S. Stepleman, ed.), North-Holland Publishing Co., Amsterdam, 1983, pp. 163–172.

23. R. S. Falk and J. E. Osborn, *Error estimates for mixed methods*, R.A.I.R.O. Numerical Analysis **14** (1980), no. 3, pp. 249–277. MR **82j:**65076

24. V. Girault and P. A. Raviart, *Finite element methods for the Navier-Stokes equations*, Springer Series in Computational Mathematics, vol. 5, Springer-Verlag, Berlin, 1986. MR **88b:**65129

25. R. Glowinski and M. F. Wheeler, *Domain decomposition and mixed finite element methods for elliptic problems*, Proceedings, First International Conference on Domain Decomposition Methods (Philadelphia) (R. Glowinski et al., eds.), SIAM, 1988, pp. 144–172. MR **90a:**65237

26. P. L. Lions, *On the Schwarz alternating method*, Proceedings, First International Conference on Domain Decomposition Methods (Philadelphia) (R. Glowinski et al., eds.), SIAM, 1988, pp. 1–42. MR **90a:**65248

27. A.M. Matsokin and S. Nepomnyaschikh, *On the Schwarz alternating method*, Preprint, Computing Center, Siberian Branch of the USSR Academy of Sciences, Novosibirsk, 1984.

28. C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal. **12** (1975), pp. 617–629. MR **52:**4595

29. W. Queck, *The convergence factor of preconditioned algorithms of the Arrow-Hurwich type*, SIAM J. Numer. Anal. **26** (1989), pp. 1016–1030. MR **90m:**65071

30. P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, Mathematical Aspects of Finite Element Methods, Lecture Notes in Mathematics, nr. 606 (I. Galligani and E. Magenes, eds.), Springer Verlag, Berlin, 1977, pp. 295–315. MR **58:**3547

31. T. Rusten and R. Winther, *A preconditioned iterative method for saddle point problems*, SIAM J. Matrix Anal. Appl. **13** (1992), pp. 887–904. MR **93a:**65043

32. ———, *Substructure preconditioners for elliptic saddle point problems*, Math. Comp. **60** (1993), pp. 23–48. MR **93d:**65108

33. D. S. Silvester and A. J. Wathen, *Fast iterative solution of stabilized Stokes systems, part II: using general block preconditioners*, SIAM J. Numer. Anal. **31** (1994), pp. 1352–1367. MR **95g:**65132

34. P. S. Vassilevski and R. D. Lazarov, *Preconditioning saddle-point problems arising from mixed finite element discretization of elliptic problems*, Report CAM 92–46, Department of Mathematics, UCLA, 1992.

35. P. S. Vassilevski and J. Wang, *An application of the abstract multilevel theory to nonconforming finite element methods*, SIAM J. Numer. Anal. **32** (1995), 235–248. CMP 95:07

36. _____, *Multilevel methods for cell–centered finite difference approximations of elliptic problems*, Preprint, 1992.

37. R. Verfürth, *A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem*, IMA J. Numerical Analysis (1984), pp. 441–455. MR **86f:**65200

38. J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review **34** (1992), pp. 581–613. MR **93k:**65029

SINTEF, P. O. Box 124 Blindern, N-0314 Oslo, Norway
E-mail address: `Torgeir.Rusten@si.sintef.no`

Center of Informatics and Computer Technology, Bulgarian Academy of Sciences, "Acad. G. Bontchev" street, Block 25 A, 1113 Sofia, Bulgaria
E-mail address: `panayot@iscbg.acad.bg`

Department of Informatics, University of Oslo, P. O. Box 1080 Blindern, N-0316 Oslo, Norway
E-mail address: `Ragnar.Winther@ifi.uio.no`