2003

A Schwarz Preconditioner for a Hybridized Mixed Method

Jay Gopalakrishnan
Portland State University, gjay@pdx.edu

A Schwarz preconditioner for a hybridized mixed method

Jayadeep Gopalakrishnan

University of Florida
Department of Mathematics, Gainesville, Florida 32611–8105.
E-mail: jayg@math.ufl.edu

*This paper is dedicated to Raytcho Lazarov, on the occasion of his 60*th *birthday.*

Abstract — In this paper, we provide a Schwarz preconditioner for the hybridized versions of the Raviart-Thomas and Brezzi-Douglas-Marini mixed methods. The preconditioner is for the linear equation for Lagrange multipliers arrived at by eliminating the flux as well as the primal variable. We also prove a condition number estimate for this equation when no preconditioner is used. Although preconditioners for the lowest order case of the Raviart-Thomas method have been constructed previously by exploiting its connection with a nonconforming method, our approach is different, in that we use a new variational characterization of the Lagrange multiplier equation. This allows us to precondition even the higher order cases of these methods.

2000 Mathematics Subject Classification: 65N55, 65N22, 65N30.

Keywords: Hybridized mixed method, Raviart-Thomas method, Brezzi-Douglas-Marini method, Schwarz, domain decomposition, preconditioner, two-level method.

1. Introduction

The subject of this paper is a Schwarz preconditioner for efficiently solving linear systems arising from the hybridized mixed method for the following Dirichlet problem:

$$-\mathrm{div}\,\boldsymbol{a}\boldsymbol{\nabla}u = f \text{ on } \Omega,$$
$$u = g \text{ on } \partial\Omega. \tag{1.1}$$

Here Ω is a polygonal domain in $\mathbb{R}^2$, $f \in L^2(\Omega)$, and $g \in H^{1/2}(\partial\Omega)$, and $\boldsymbol{a}(\boldsymbol{x})$ is a symmetric 2×2 matrix function of $\boldsymbol{x} \in \Omega$ that is uniformly positive definite and bounded in Ω.

Efficient solution strategies for mixed methods have been investigated earlier by many authors [3, 4, 6, 7, 11, 20, 22, 24] using a wide variety of techniques: V-cycle and W-cycle methods were given in [4] and [3]. An equivalence of the mixed method with a nonconforming method was utilized in [6]. In [24], it was shown that it suffices to precondition a spectrally equivalent discontinuous Galerkin like bilinear form. All these works dealt with the non-hybridized form of the mixed method. In contrast, in this paper we consider the hybridized version of a mixed method. This paper also differs from other works that have dealt with solution strategies after hybridization in the context of substructuring, notably [20]. The situation we have in mind is one where hybridization is done at the element level, rather than in a nonoverlapping domain decomposition method.

There are a few earlier works on preconditioning the hybridized form of the mixed method that we should note here. A balancing domain decomposition method for the hybridized mixed method is discussed in [14]. More results on domain decomposition algorithms utilizing the hybridization concepts can be found in [25]. While these works considered systems that couple the so called Lagrange multiplier unknowns together with the primal variable, we consider a system that involves only the Lagrange multiplier unknowns.

To precisely describe this system, we start by describing a hybridized mixed method. The method is obtained by using Lagrange multipliers to enforce continuity constraints of a vector finite element space in a standard mixed method. For the sake of definiteness let us consider the hybridized version of the Raviart-Thomas (RT) mixed method [23]. As we shall see, our considerations hold if the Brezzi-Douglas-Marini (BDM) mixed method [9] is used instead. On any triangle τ, let $P_d(\tau)$ denote the set of polynomials (in the variable $\boldsymbol{x} \in \mathbb{R}^2$) of degree at most d on τ, and let $\boldsymbol{R}_d(\tau) = \boldsymbol{x}P_d(\tau) + (P_d(\tau))^2$. Let $\mathcal{T}_h$ be a triangulation of Ω, and $\mathcal{E}_h$ be the set of its interior edges. Let $\boldsymbol{c}(x) = \boldsymbol{a}(x)^{-1}$, and $\langle \cdot, \cdot \rangle_Z$ for any space Z denote the duality pairing in Z. Define spaces

$$\boldsymbol{R}_h = \{\boldsymbol{r} : \boldsymbol{r}|_\tau \in \boldsymbol{R}_d(\tau) \text{ for all } \tau \in \mathcal{T}_h\},$$
$$T_h = \{p : p|_\tau \in P_d(\tau) \text{ for all } \tau \in \mathcal{T}_h\},$$
$$S_h = \{\lambda : \lambda|_e \in P_d(e) \text{ for all } e \in \mathcal{E}_h\},$$

and operators $\boldsymbol{\mathcal{A}} : \boldsymbol{R}_h \mapsto \boldsymbol{R}'_h$, $\mathcal{B} : \boldsymbol{R}_h \mapsto T'_h$, and $\mathcal{C} : \boldsymbol{R}_h \mapsto S'_h$ by

$$\langle \boldsymbol{\mathcal{A}}\boldsymbol{q}, \boldsymbol{r} \rangle_{\boldsymbol{R}_h} = \int_\Omega \boldsymbol{c}\,\boldsymbol{q} \cdot \boldsymbol{r}\, dx$$

$$\langle \mathcal{B}\boldsymbol{q}, p \rangle_{T_h} = \sum_{\tau \in \mathcal{T}_h} - \int_\tau p\, \text{div}\, \boldsymbol{q}\, dx$$

$$\langle \mathcal{C}\boldsymbol{q}, \mu \rangle_{S_h} = \sum_{e \in \mathcal{E}_h} \int_e \mu\, [\![\boldsymbol{q}]\!]\, ds.$$

Here, for any edge $e \in \mathcal{E}_h$, if $\tau_+, \tau_- \in \mathcal{T}_h$ are the triangles that share edge e with outward normals $\boldsymbol{n}_+$ and $\boldsymbol{n}_-$ respectively, then $[\![\boldsymbol{q}]\!]$ on e equals $(\boldsymbol{q}|_{\tau_+} \cdot \boldsymbol{n}_+) + (\boldsymbol{q}|_{\tau_-} \cdot \boldsymbol{n}_-)$.

The hybridized mixed method using the above spaces defines an approximate solution triple $(\boldsymbol{q}_h, u_h, \lambda_h) \in \boldsymbol{R}_h \times T_h \times S_h$ as the unique solution of

$$\begin{pmatrix} \boldsymbol{\mathcal{A}} & \mathcal{B}^t & \mathcal{C}^t \\ \mathcal{B} & 0 & 0 \\ \mathcal{C} & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{q}_h \\ u_h \\ \lambda_h \end{pmatrix} = \begin{pmatrix} \mathcal{G} \\ \mathcal{F} \\ 0 \end{pmatrix}, \tag{1.2}$$

where $\mathcal{G}$ and $\mathcal{F}$ are functionals on $\boldsymbol{R}_h$ and T_h, respectively, given by

$$\langle \mathcal{G}, \boldsymbol{r} \rangle_{\boldsymbol{R}_h} = - \int_{\partial\Omega} g(\boldsymbol{r} \cdot \boldsymbol{n})\, ds,$$

$$\langle \mathcal{F}, p \rangle_{T_h} = - \int_\Omega fp\, dx.$$

As is well known [10], the variables $\boldsymbol{q}_h$ and u_h can be eliminated from (1.2) to yield an equation involving just the multiplier λ_h:

$$\begin{aligned} (\mathcal{C}\boldsymbol{\mathcal{A}}^{-1}\mathcal{C}^t &- \mathcal{C}\boldsymbol{\mathcal{A}}^{-1}\mathcal{B}^t(\mathcal{B}\boldsymbol{\mathcal{A}}^{-1}\mathcal{B}^t)^{-1}\mathcal{B}\boldsymbol{\mathcal{A}}^{-1}\mathcal{C}^t)\lambda_h \\ &= -\mathcal{C}\boldsymbol{\mathcal{A}}^{-1}\mathcal{B}^t(\mathcal{B}\boldsymbol{\mathcal{A}}^{-1}\mathcal{B}^t)^{-1}(\mathcal{B}\boldsymbol{\mathcal{A}}^{-1}\mathcal{G} - \mathcal{F}) + \mathcal{C}\boldsymbol{\mathcal{A}}^{-1}\mathcal{G}. \end{aligned} \tag{1.3}$$

There are many reasons why one should design an implementation of the mixed method that first solves (1.3). First of all, (1.3) can easily be proved to be a symmetric positive definite system for λ_h. Therefore, it is more suited for modern iterative solution methods (like the conjugate gradient method) compared to the indefinite system (1.2). Moreover, the number of unknowns in (1.3) is clearly much less than that of (1.2). Yet another reason is that once λ_h is computed, the other components of the solution triple, namely $\boldsymbol{q}_h$ and u_h can be computed inexpensively in a completely local fashion [10] (element by element). Finally, let us also note that implementing (1.3) is preferred over implementing the non-hybridized mixed method, because the former yields the Lagrange multiplier λ_h which can be used to arrive at a locally post-processed solution of enhanced accuracy, as shown in [2]. Therefore, it is of considerable practical interest to design efficient solution methods for solving (1.3).

In this paper, we will construct a Schwarz preconditioner for efficiently solving (1.3). In the next section we will show that (1.3), without any preconditioner, gives rise to badly conditioned systems for small mesh sizes. When the Schwarz preconditioner is used, the preconditioned system is uniformly well conditioned.

Schwarz preconditioners, sometimes known as overlapping domain decomposition preconditioners, have been adapted to various applications ever since the early works of [16, 17] showed its suitability for some standard applications. In adapting it to precondition (1.3), one of the difficulties that we are faced with is that the multiplier spaces on refinements of a mesh are not nested. In this paper we will overcome this difficulty in the context of an "additive two-level" method, by introducing an intergrid transfer (or prolongation) operator (see Section 3, and further examples in Section 5). We use a strategy for analysis similar to that in [24], which in turn is based on techniques introduced in [16, 17].

Another difficulty is that the spectral nature of the operator in (1.3) is not obvious. It is perhaps this difficulty that has thus far prevented the design of preconditioners for the hybridized mixed method in the higher order case. In the case of the lowest order hybridized RT method, it is possible to conclude from [2] that (1.3) is equivalent to a system arising from the P_1-nonconforming method. Then, it suffices to precondition the latter. This has been exploited in earlier papers [6, 12, 19]. Nonetheless, the nature of the left hand side of (1.3) in the higher order case remained unclear. However, we can now overcome this difficulty because of a recently developed variational characterization of (1.3). We will briefly review this characterization in the next section.

2. Equation for the Lagrange multiplier

In this section we investigate the equation determining the Lagrange multiplier, namely (1.3), further. We will recall a recently developed variational characterization of (1.3) in terms of certain lifting maps, provide a norm equivalence for the resulting bilinear form, and prove a condition number estimate for (1.3).

Suppose we are given a nodal basis for S_h, say $\{\eta_i\}_{i=1}^M$, such that each η_i is supported on a single edge of $\mathcal{E}_h$. For example, η_i is one of the first $d+1$ Legendre polynomials on one edge and is zero on all the other edges. Equation (1.3) then yields a matrix equation for the vector of coefficients of λ_h in the $\{\eta_i\}$-basis, which we denote by Λ:

$$\mathrm{E}\Lambda = \mathrm{b}. \tag{2.1}$$

Obviously, the $M \times M$ matrix E can be computed once the matrices of the operators $\boldsymbol{\mathcal{A}}$, $\mathcal{B}$, and $\mathcal{C}$ are computed. However, there is an easier way to compute E. It turns out to be

the stiffness matrix of a mesh dependent bilinear form $a_h(\cdot,\cdot)$ defined below. After defining $a_h(\cdot,\cdot)$, we will estimate the condition number of E.

Define lifting operators $\boldsymbol{Q} : S_h \mapsto \boldsymbol{R}_h$ and $\mathcal{U} : S_h \mapsto T_h$ element by element: On each $\tau \in \mathcal{T}_h$, $(\boldsymbol{Q}\lambda)|_\tau$ and $(\mathcal{U}\lambda)|_\tau$ are defined by

$$\int_\tau \boldsymbol{c}\,(\boldsymbol{Q}\lambda) \cdot \boldsymbol{q}\,dx - \int_\tau (\mathcal{U}\lambda)\,\mathrm{div}\,\boldsymbol{q}\,dx = -\int_{\partial\tau\setminus\partial\Omega} \lambda \boldsymbol{q} \cdot \boldsymbol{n}_\tau\,ds \tag{2.2}$$

$$\int_\tau p\,\mathrm{div}(\boldsymbol{Q}\lambda)\,dx = 0 \tag{2.3}$$

for all $\boldsymbol{q} \in \boldsymbol{R}_d(\tau)$ and $p \in P_d(\tau)$, where $\boldsymbol{n}_\tau$ denotes the outward unit normal vector on $\partial\tau$. Let

$$a_h(\lambda,\mu) = \int_\Omega c\,(\boldsymbol{Q}\lambda) \cdot (\boldsymbol{Q}\mu)\,dx.$$

The following theorem shows that the nature of the discrete linear system (2.1) that determine λ_h is intimately related to the nature of the bilinear form $a_h(\cdot,\cdot)$. A proof can be found in [13].

Theorem 2.1. *The Lagrange multiplier component of the hybridized mixed method solution, namely λ_h, is the unique element of S_h that satisfies*

$$a_h(\lambda_h,\mu) = \int_\Omega f\,\mathcal{U}\mu + \int_{\partial\Omega} g\,\boldsymbol{Q}\mu \cdot \boldsymbol{n} \quad \text{for all } \mu \in S_h, \tag{2.4}$$

where $\boldsymbol{n}$ denotes the outward unit normal vector on $\partial\Omega$. Moreover, $\mathrm{E}_{ij} = a_h(\eta_j,\eta_i)$ and $\mathrm{b}_i = \int_\Omega f\,\mathcal{U}\eta_i + \int_{\partial\Omega} g\,\boldsymbol{Q}\eta_i \cdot \boldsymbol{n}$, for all $i,j = 1,2,\ldots,M$.

In particular, it follows from this theorem that E is a sparse matrix. Indeed, the liftings of η_i are supported only on the two triangles that share the edge which forms the support of η_i. In the lowest order case, this means that the matrix E has *at most four* nonzero off-diagonal entries. In the general case, E is a matrix of $(d+1) \times (d+1)$ blocks with at most four off-diagonal blocks in each block column.

At this point, let us note that the definition of $a_h(\cdot,\cdot)$ depends only on the divergence free members of the RT-space. Specifically, the lifting operator $\boldsymbol{Q}$ in the definition of $a_h(\cdot,\cdot)$ can be given solely using $\boldsymbol{R}_d^0(\tau) = \{\boldsymbol{r} \in \boldsymbol{R}_d(\tau) : \mathrm{div}\,\boldsymbol{r} = 0\}$ as follows: $(\boldsymbol{Q}\mu)|_\tau$ for every $\tau \in \mathcal{T}_h$ is the unique element of $\boldsymbol{R}_d^0(\tau)$ satisfying

$$\int_\tau \boldsymbol{c}\,\boldsymbol{Q}\mu \cdot \boldsymbol{r}\,dx = -\int_{\partial\tau\setminus\partial\Omega} \mu\,\boldsymbol{r} \cdot \boldsymbol{n}_\tau\,ds \quad \text{for all } \boldsymbol{r} \in \boldsymbol{R}_d^0(\tau).$$

It is well known that the divergence free subspaces of the vector BDM space and the vector RT space (which we denoted by $\boldsymbol{R}_d^0(\tau)$) on one triangle coincide. This means that the bilinear form $a_h(\cdot,\cdot)$ and the left hand side matrix in (2.1) are identical to the corresponding ones arising in hybridization of the BDM-method. Therefore, for the purposes of preconditioning the Lagrange multiplier equation, we can ignore the differences between the BDM and RT methods.

We now clarify the nature of the norm generated by $a_h(\cdot,\cdot)$. For any domain D we denote by $\|\cdot\|_{L^2(D)}$ the $L^2(D)$-norm (or the $(L^2(D))^2$-norm, as appropriate). We identify

$\lambda \in S_h$ with its extension by zero to edges on $\partial\Omega$ for simplifying notation, so that e.g., $\int_{\partial\tau \setminus \partial\Omega} \lambda \, ds = \int_{\partial\tau} \lambda \, ds$. Define

$$m_\tau(\lambda) = \frac{1}{|\partial\tau|} \int_{\partial\tau} \lambda,$$

$$\|\lambda\|_{h,D} = \left(\sum_{\tau \in \mathcal{T}_h, \, \tau \subseteq \bar{D}} \|\lambda - m_\tau(\lambda)\|_{L^2(\partial\tau)}^2 \frac{1}{|\partial\tau|} \right)^{1/2}, \quad \text{and}$$

$$\|\lambda\|_{h,D} = \left(\sum_{\tau \in \mathcal{T}_h, \, \tau \subseteq \bar{D}} \|\lambda\|_{L^2(\partial\tau)}^2 |\partial\tau| \right)^{1/2}.$$

When the domain under consideration is Ω we use $\|\cdot\|_h$ and $\|\cdot\|_h$ to denote $\|\cdot\|_{h,\Omega}$ and $\|\cdot\|_{h,\Omega}$ respectively. The following theorem shows that the norm generated by $a_h(\cdot,\cdot)$ is equivalent to the more transparent norm $\|\cdot\|_h$.

Theorem 2.2. *For any triangle K, there are positive constants C_1 and C_2 depending only on d, $\boldsymbol{c}$, and the minimal angle of K, such that*

$$C_1 \|\lambda\|_{h,K}^2 \leqslant \int_K \boldsymbol{c} \, |\boldsymbol{\mathfrak{Q}}\lambda|^2 \, dx \leqslant C_2 \|\lambda\|_{h,K}^2, \quad \text{for all } \lambda \in S_h.$$

Before we prove this theorem, we state one other theorem that we will prove in this section. Although the main use of Theorem 2.2 is in the analysis of our Schwarz preconditioner, as one of its immediate applications, we can estimate the condition number of the stiffness matrix E when no preconditioner is used. It is generally accepted that the condition number of (1.3) must be $O(h^{-2})$ on a quasiuniform mesh on mesh size h, although we have not been able to locate a precise statement to this effect in the literature. The following heuristic argument is often given: Since the Lagrange multipliers approximate the exact solution on the edges of the mesh [2], Equation (1.3) should be a discretization of the elliptic second order equation in (1.1), and hence should exhibit the same growth in condition number that other discretizations suffer. We give a precise bound in the following theorem. We adopt the convention of denoting by C (with or without subscripts) a generic constant independent of h. In general, its value differs at different occurrences.

Theorem 2.3. *Suppose $\mathcal{T}_h$ is a quasiuniform mesh of mesh size h. Then, there are positive constants C_3 and C_4 independent of h such that*

$$C_3 \|\lambda\|_h^2 \leqslant a_h(\lambda, \lambda) \leqslant C_4 h^{-2} \|\lambda\|_h^2 \quad \text{for all } \lambda \in S_h. \tag{2.5}$$

Consequently, the spectral condition number of E in (2.1) is $O(h^{-2})$.

In the remainder of this section we prove Theorems 2.2 and 2.3. The proof of Theorem 2.2 is based on the following lemma.

Lemma 2.1. *The function $\boldsymbol{\mathfrak{Q}}\lambda$ is zero on $K \in \mathcal{T}$ if and only if λ is constant on ∂K.*

Proof. From the definition of $\boldsymbol{\mathfrak{Q}}\lambda$, note that

$$\int_K \boldsymbol{c} \, (\boldsymbol{\mathfrak{Q}}\lambda) \cdot (\boldsymbol{\mathfrak{Q}}\lambda) \, dx = - \int_{\partial K} \lambda \, (\boldsymbol{\mathfrak{Q}}\lambda) \cdot \boldsymbol{n}_K \, ds. \tag{2.6}$$

By integration by parts, the right hand side above equals $-\lambda \int_K \operatorname{div}(\boldsymbol{\mathfrak{Q}}\lambda)\, dx$, whenever λ is constant on ∂K. Since $\operatorname{div}(\boldsymbol{\mathfrak{Q}}\lambda|_K) = 0$, the right hand side vanishes, so $\boldsymbol{\mathfrak{Q}}\lambda$ is zero on K.

Conversely, suppose $\boldsymbol{\mathfrak{Q}}\lambda$ is zero on K. Then (2.2) becomes

$$\int_K (\mathfrak{U}\lambda)\operatorname{div}\boldsymbol{r}\, dx = \int_{\partial K} \lambda\, \boldsymbol{r}\cdot\boldsymbol{n}_K\, ds, \quad \text{for all } \boldsymbol{r} \in \boldsymbol{R}_d(K),$$

After integrating by parts,

$$-\int_K \boldsymbol{r}\cdot\boldsymbol{\nabla}(\mathfrak{U}\lambda)\, dx + \int_{\partial K}(\boldsymbol{r}\cdot\boldsymbol{n}_K)\mathfrak{U}\lambda\, ds = \int_{\partial K} \lambda\,\boldsymbol{r}\cdot\boldsymbol{n}_K\, ds, \tag{2.7}$$

for all $\boldsymbol{r} \in \boldsymbol{R}_d(K)$. In this equation, we can choose $\boldsymbol{r}$ such that

$$\int_K \boldsymbol{r}\cdot\boldsymbol{p}_{d-1}\, dx = 0, \quad \text{for all } \boldsymbol{p}_{d-1} \in P_{d-1}(K)^2, \text{ and}$$

$$\boldsymbol{r}\cdot\boldsymbol{n}_K = \lambda - (\mathfrak{U}\lambda)|_{\partial K}, \quad \text{on } \partial K.$$

Then, (2.7) gives

$$\int_{\partial K}(\lambda - \mathfrak{U}\lambda)^2\, ds = 0,$$

so $\mathfrak{U}\lambda$ coincides with λ on the boundary ∂K. Resorting to (2.7) again, and using the fact $(\mathfrak{U}\lambda - \lambda)|_{\partial K} = 0$, we find that

$$\int_K \boldsymbol{r}\cdot\boldsymbol{\nabla}(\mathfrak{U}\lambda)\, dx = 0, \quad \text{for all } \boldsymbol{r} \in \boldsymbol{R}_d(\tau),$$

so $\mathfrak{U}\lambda$ is constant on K. Since λ coincides with $\mathfrak{U}\lambda$ on ∂K, this implies that λ is constant on ∂K. $\square$

We can now prove Theorem 2.2 using this lemma.

Proof. (*Proof of Theorem 2.2.*) Let us first prove the upper bound of the theorem. From (2.6), we have

$$\int_K \boldsymbol{c}\,|\boldsymbol{\mathfrak{Q}}\lambda|^2\, dx = -\int_{\partial K}(\lambda - m_K(\lambda))(\boldsymbol{n}_K\cdot\boldsymbol{\mathfrak{Q}}\lambda)\, ds$$

It follows by a scaling argument using the Piola map, a trace theorem on a fixed reference triangle, and Cauchy-Schwarz inequality that

$$\int_K \boldsymbol{c}\,|\boldsymbol{\mathfrak{Q}}\lambda|^2\, dx \leqslant C|\partial K|^{-1/2}\|\lambda - m_K(\lambda)\|_{L^2(\partial K)}\|\boldsymbol{\mathfrak{Q}}\lambda\|_{L^2(K)},$$

thus proving the upper bound.

To prove the lower bound, we use Lemma 2.1, and a scaling argument. Let $\hat{K}$ denote a fixed reference triangle. For any symmetric positive definite 2×2 matrix valued function $\boldsymbol{\alpha}(\hat{\boldsymbol{x}})$ on $\hat{K}$, define liftings $\widehat{\boldsymbol{\mathfrak{Q}}}_{\boldsymbol{\alpha}}\hat{\lambda} \in \boldsymbol{R}_d(\hat{K})$ and $\widehat{\mathfrak{U}}_{\boldsymbol{\alpha}}\hat{\lambda} \in P_d(\hat{K})$ on $\hat{K}$ by

$$\int_{\hat{K}} \boldsymbol{\alpha}\,(\widehat{\boldsymbol{\mathfrak{Q}}}_{\boldsymbol{\alpha}}\hat{\lambda})\cdot\boldsymbol{r}\, dx - \int_{\hat{K}}(\widehat{\mathfrak{U}}_{\boldsymbol{\alpha}}\hat{\lambda})\operatorname{div}\boldsymbol{r}\, dx = -\int_{\partial\hat{K}}\hat{\lambda}\,\boldsymbol{r}\cdot\boldsymbol{n}_{\hat{K}}\, ds,$$

$$\int_{\hat{K}} v\operatorname{div}(\widehat{\boldsymbol{\mathfrak{Q}}}_{\boldsymbol{\alpha}}\hat{\lambda})\, dx = 0,$$

for all $\boldsymbol{r} \in \boldsymbol{R}_d(\hat{K})$ and $v \in P_d(\hat{K})$. By Lemma 2.1, $\widehat{\mathbf{Q}}_{\boldsymbol{\alpha}}\hat{\lambda} = 0$ implies that $\hat{\lambda}$ is constant on $\partial\hat{K}$. Therefore,

$$\|\widehat{\mathbf{Q}}_{\boldsymbol{\alpha}}\hat{\lambda}\|_{L^2(\hat{K})} \geqslant \hat{C}(\boldsymbol{\alpha}) \inf_{\kappa \in \mathbb{R}} \|\hat{\lambda} - \kappa\|_{L^2(\partial\hat{K})}, \tag{2.8}$$

for some constant $\hat{C}(\boldsymbol{\alpha})$ independent of $\hat{\lambda}$. We now relate the liftings on K with these liftings on the reference element. Let $\hat{\boldsymbol{x}} \mapsto \boldsymbol{x} \equiv D_K\hat{\boldsymbol{x}} + \boldsymbol{d}_K$ be the affine isomorphism that maps $\hat{K}$ one-one onto K. For scalar valued functions $\mu(\boldsymbol{x})$, we define $\hat{\mu}(\hat{\boldsymbol{x}}) = \mu(\boldsymbol{x})$, while for vector valued functions $\boldsymbol{r}(\boldsymbol{x})$, we define $\hat{\boldsymbol{r}}(\hat{\boldsymbol{x}}) = |\det D_K|D_K^{-1}\boldsymbol{r}(\boldsymbol{x})$. Then, it is easily seen that if we set $\hat{\boldsymbol{c}}(\hat{\boldsymbol{x}}) = |\det D_K|^{-1}D_K^t\boldsymbol{c}(\boldsymbol{x})D_K$, we have

$$\widehat{\mathbf{Q}}_{\hat{\boldsymbol{c}}}\hat{\lambda} = \widehat{\mathbf{Q}\lambda}. \tag{2.9}$$

In view of (2.8) and (2.9), we get by a scaling argument that

$$\begin{aligned} C\|\mathbf{Q}\lambda\|_{L^2(K)} &\geqslant \|\widehat{\mathbf{Q}\lambda}\|_{L^2(\hat{K})} = \|\widehat{\mathbf{Q}}_{\hat{\boldsymbol{c}}}\hat{\lambda}\|_{L^2(\hat{K})} \\ &\geqslant \hat{C}(\hat{\boldsymbol{c}}) \inf_{\kappa \in \mathbb{R}} \|\hat{\lambda} - \kappa\|_{L^2(\partial\hat{K})}. \end{aligned}$$

Mapping back, we have

$$C\|\mathbf{Q}\lambda\|_{L^2(K)}^2 \geqslant \frac{\hat{C}(\hat{\boldsymbol{c}})^2}{|\partial K|} \inf_{\kappa \in \mathbb{R}} \|\lambda - \kappa\|_{L^2(\partial K)}^2 = \frac{\hat{C}(\hat{\boldsymbol{c}})^2}{|\partial K|} \|\lambda - m_K(\lambda)\|_{L^2(\partial K)}^2.$$

Thus, the estimate of the theorem will follow provided we can show that $\hat{C}(\hat{\boldsymbol{c}})$ is bounded uniformly away from zero. It is easily seen that we can choose $\hat{C}(\cdot)$ to be a positive continuous function. Moreover, all $\hat{\boldsymbol{c}}$ obtained by transforming $\boldsymbol{c}(\boldsymbol{x})$, lie in a compact set $\{\boldsymbol{\alpha} : C_5 \leqslant \|\boldsymbol{\alpha}\|_{\ell^2} \leqslant C_6\}$, because $\boldsymbol{a}(\boldsymbol{x})$ is uniformly positive definite and bounded on Ω. Taking the minimum of the function $\hat{C}(\boldsymbol{\alpha})$ over this compact set, we have the required result. $\qquad\square$

Next, we prove Theorem 2.3. We shall use the inverse estimate

$$\|\lambda\|_{h,\tau} \leqslant \frac{2}{|\partial\tau|}\|\lambda\|_{h,\tau}, \tag{2.10}$$

which immediately follows from the definition of our norms:

$$\begin{aligned} \|\lambda\|_{h,\tau}^2 &= \frac{1}{|\partial\tau|}\|\lambda - m_\tau(\lambda)\|_{L^2(\partial\tau)}^2 \\ &\leqslant \frac{2}{|\partial\tau|}\left(\|\lambda\|_{L^2(\partial\tau)}^2 + \left(\frac{1}{|\partial\tau|}\int_{\partial\tau}\lambda\right)^2|\partial\tau|\right) \leqslant \frac{4}{|\partial\tau|}\|\lambda\|_{L^2(\partial\tau)}^2. \end{aligned}$$

Proof. (*Proof of Theorem 2.3.*) The upper bound is a direct consequence of (2.10) and Theorem 2.2. To prove the lower bound, define u_λ for any $\lambda \in S_h$ element by element as follows: On any $\tau \in \mathcal{T}_h$, $u_\lambda|_\tau \in P_{d+1}(\tau)$ satisfies

$$\int_\tau \boldsymbol{\nabla}u_\lambda \cdot \boldsymbol{\nabla}v\,dx + \frac{1}{|\partial\tau|}\int_{\partial\tau}u_\lambda v\,ds = \frac{1}{|\partial\tau|}\int_{\partial\tau}\lambda v\,ds \quad \text{for all } v \in P_{d+1}(\tau). \tag{2.11}$$

Obviously, given any $\lambda \in S_h$, such a u_λ is uniquely defined. In the remainder of this proof we show that

$$\|\lambda\|_h \leqslant C\|u_\lambda\|_{L^2(\Omega)} \leqslant C\|\lambda\|_h \quad \text{for all } \lambda \in S_h. \tag{2.12}$$

Clearly this will prove the required lower bound.

The first inequality of (2.12) follows easily from a local scaling argument, so we will only prove that $\|u_\lambda\|_{L^2(\Omega)} \leqslant C\|\lambda\|_h$. Choosing $v \equiv 1$ in (2.11), we have

$$\int_{\partial\tau} (u_\lambda - \lambda)\, ds = 0.$$

This implies that

$$
\begin{aligned}
\frac{1}{|\partial\tau|} \int_{\partial\tau} (u_\lambda - \lambda)^2\, ds &= -\int_\tau |\boldsymbol{\nabla} u_\lambda|^2\, dx + \frac{1}{|\partial\tau|} \int_{\partial\tau} \lambda(u_\lambda - \lambda)\, ds \\
&= -\int_\tau |\boldsymbol{\nabla} u_\lambda|^2\, dx + \frac{1}{|\partial\tau|} \int_{\partial\tau} (\lambda - m_\tau(\lambda))(u_\lambda - \lambda)\, ds \\
&\leqslant \frac{1}{|\partial\tau|} \int_{\partial\tau} (\lambda - m_\tau(\lambda))(u_\lambda - \lambda)\, ds.
\end{aligned}
$$

Therefore,

$$\frac{1}{|\partial\tau|}\|u_\lambda - \lambda\|_{L^2(\partial\tau)}^2 \leqslant \|\lambda\|_{h,\tau}^2. \tag{2.13}$$

Now, it is readily verified that when λ is constant on $\partial\tau$, u_λ is constant on τ, so a scaling argument shows that

$$\|\boldsymbol{\nabla} u_\lambda\|_{L^2(\tau)}^2 \leqslant C\|\lambda\|_{h,\tau}^2. \tag{2.14}$$

Moreover, on an interior edge e shared by two triangles $\tau_+, \tau_- \in \mathcal{T}_h$, the jump of u_λ across e, denoted by $[u_\lambda]$, satisfies

$$\frac{1}{|e|}\|\,[u_\lambda]\,\|_{L^2(e)}^2 = \frac{1}{|e|}\|(u_\lambda|_{\tau_+} - \lambda) - (u_\lambda|_{\tau_-} - \lambda)\|_{L^2(e)}^2 \leqslant C\|\lambda\|_{h,\tau_+\cup\tau_-}^2,$$

because of (2.13). Thus,

$$\left(\sum_{\tau\in\mathcal{T}_h} \|\boldsymbol{\nabla} u_\lambda\|_{L^2(\tau)}^2 + \sum_{e\in\mathcal{E}_h} \frac{1}{|e|}\|\,[u_\lambda]\,\|_{L^2(e)}^2 \right)^{1/2} \leqslant C\|\lambda\|_h.$$

The left hand side of the inequality above defines a norm previously studied [24]. In particular, the following Poincaré inequality is well known (see [24, Theorem 3.1], [18], [1, Lemma 2.1], or [8]):

$$C\|u_\lambda\|_{L^2(\Omega)}^2 \leqslant \sum_{\tau\in\mathcal{T}_h} \|\boldsymbol{\nabla} u_\lambda\|_{L^2(\tau)}^2 + \sum_{e\in\mathcal{E}_h} \frac{1}{|e|}\|\,[u_\lambda]\,\|_{L^2(e)}^2.$$

This, together with the estimates above, yields (2.12), and the lower bound of the theorem follows.

The assertion on the spectral condition number of E follows from (2.5): Expand $\lambda \in S_h$ in the basis $\{\eta_i\}$ as

$$\lambda = \sum_{i=1}^M \ell_i\, \eta_i.$$

The vector of coefficients $\ell = (\ell_1, \ell_2, \ldots, \ell_M)^t \in \mathbb{R}^M$ satisfies $C(\ell \cdot \ell)h^2 \leqslant \|\lambda\|_h^2 \leqslant C(\ell \cdot \ell)h^2$, because the basis functions η_i are local. Consequently, $Ch^2(\ell \cdot \ell) \leqslant \mathrm{E}\ell \cdot \ell \leqslant C(\ell \cdot \ell)$ for all $\ell \in \mathbb{R}^M$. $\qquad\square$

3. A Schwarz preconditioner

A basic assumption in Schwarz algorithms [16, 17] is that the mesh wherein solution is sought, namely $\mathcal{T}_h$, is a refinement of a coarser mesh, say $\mathcal{T}_H$, consisting of coarse elements $\{\Omega_i\}_{i=1}^N$. We assume for the purposes of analysis that $\mathcal{T}_h$ is quasiuniform of mesh size h, and that $\mathcal{T}_H$ is quasiuniform of mesh size H ($H > h$). Schwarz algorithms use solutions on overlapping subdomains $\widetilde{\Omega}_i$, $i = 1, \ldots, N$, that cover Ω. The closure of each $\widetilde{\Omega}_i$ is a union of triangles of $\mathcal{T}_h$, and contains the coarse triangle Ω_i. (By definition, Ω and $\widetilde{\Omega}_i$ are open sets.) We assume that there are fixed numbers δ and ρ such that

$$\text{dist}\,(\partial\widetilde{\Omega}_i \cap \Omega, \partial\Omega_i \cap \Omega) \geqslant \delta H \quad \text{for all } i = 1, \ldots N, \quad \text{and} \tag{3.1}$$

$$\text{every point of } \Omega \text{ is in at most } \rho \text{ subdomains in } \{\widetilde{\Omega}_i\}_{i=1}^N. \tag{3.2}$$

The preconditioner we describe in this section uses solutions on the coarse mesh $\mathcal{T}_H$ using the lowest order space of multipliers S_H^0. The generally accepted intuitive reason for using a coarse space is that it helps global propagation of information during an iterative solution process. It is therefore intuitive to set a coarse space based on the lowest order space S_H^0, although the fine space S_h is not in general of lowest order.

The main difficulty in incorporating information from coarse solutions into the precon-ditioner arises from the fact that $S_H^0 \not\subseteq S_h$. We overcome this by introducing intergrid transfer operators $I_h : S_H^0 \mapsto S_h$. To this end, we exploit a relationship between the P1-nonconforming space V_h^0 and the lowest order space of multipliers S_h^0. Recall that $S_h^0 = \{ \lambda : \lambda|_e$ is a constant for every $e \in \mathcal{E}_h\}$, and $V_h^0 = \{ v : v|_\tau$ is linear for all $\tau \in \mathcal{T}_h$, v is continuous at midpoints of all $e \in \mathcal{E}_h$ and zero at midpoints of edges $e \subseteq \partial\Omega\}$. We establish an isomorphism between these spaces, namely $X_h : S_h^0 \mapsto V_h^0$, by

$$(X_h\lambda)(x_e) = \lambda|_e,$$

where x_e denotes the midpoint of edge e. Obviously $X_h^{-1} : V_h^0 \mapsto S_h^0$ is well-defined, and actions of both X_h and X_h^{-1} are easily implementable in computations. Define a seminorm $|\cdot|_{H^1(\mathcal{T}_h)}$ by

$$|w|_{H^1(\mathcal{T}_h)}^2 = \sum_{\tau \in \mathcal{T}_h} \|\boldsymbol{\nabla} w\|_{L^2(\tau)}^2.$$

The following lemma is established easily by means of scaling arguments, so we omit its proof.

Lemma 3.1. *For all $w \in V_h^0$ and $\lambda \in S_h^0$,*

$$\|X_h^{-1}w\|_h \leqslant C|w|_{H^1(\mathcal{T}_h)},$$
$$\|X_h^{-1}w\|_h \leqslant C\|w\|_{L^2(\Omega)},$$
$$|X_h\lambda|_{H^1(\mathcal{T}_h)} \leqslant C\|\lambda\|_h^2.$$

We can now define an intergrid transfer operator $I_h : S_H^0 \mapsto S_h$ by

$$I_h = X_h^{-1}I_h^V X_H, \tag{3.3}$$

where $I_h^V : V_H^0 \mapsto V_h^0$ is an intergrid transfer operator for P1-nonconforming spaces (see Fig 1) defined as follows: Let z_i, $i = 1, \ldots M_H$, denote the interior vertices of $\mathcal{T}_H$, and let
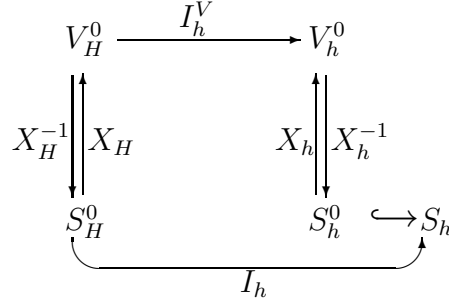
Figure 1. Operators connecting various spaces

$\tau(z_i)$ be one of the triangles of $\mathcal{T}_H$ which has z_i as a vertex. For every $w \in V_H^0$, $I_h^V w$ is the continuous function that is linear on every $\tau \in \mathcal{T}_H$, vanishes on $\partial\Omega$, and satisfies

$$(I_h^V w)(z_i) = \lim_{\substack{z \to z_i \\ z \in \tau(z_i)}} w(z). \tag{3.4}$$

In general w can have different limiting values at a vertex z_i, depending on from which triangle we approach z_i. We have set $I_h^V w$ at z_i to equal (any) one of such values. Note that $I_h^V w$ is a continuous function that is linear on each triangle of the fine mesh $\mathcal{T}_h$. In particular, it is in V_h^0. (We will give more examples of intergrid transfer operators in Section 5.)

Let $a_H(\cdot, \cdot)$ be defined analogously to $a_h(\cdot, \cdot)$, but with respect to $\mathcal{T}_H$. Define $S_i = \{\lambda \in S_h :$ support of λ is contained in $\widetilde{\Omega}_i\}$. We denote by $\langle\cdot,\cdot\rangle_h$, $\langle\cdot,\cdot\rangle_H$, and $\langle\cdot,\cdot\rangle_i$ the duality pairing in spaces S_h, S_H^0 and S_i repectively. We identify functions in S_i by their extension by zero, so we will often use $\langle\cdot,\cdot\rangle_h$ for $\langle\cdot,\cdot\rangle_i$. Denoting dual spaces by primes, let the operators $A_h : S_h \mapsto S_h'$, $A_H : S_H^0 \mapsto (S_H^0)'$, $A_i : S_i \mapsto S_i'$, $Q_i : S_h' \mapsto S_i'$, and $\widetilde{Q}_H : S_h' \mapsto (S_H^0)'$ be defined by

$$
\begin{aligned}
\langle A_h \lambda, \mu \rangle_h &= a_h(\lambda, \mu) &&\text{for all } \lambda, \mu \in S_h, \\
\langle A_H \lambda, \mu \rangle_H &= a_H(\lambda, \mu) &&\text{for all } \lambda, \mu \in S_H^0, \\
\langle A_i \lambda, \mu \rangle_h &= a_h(\lambda, \mu) &&\text{for all } \lambda, \mu \in S_i, \\
\langle Q_i \alpha, \mu \rangle_h &= \langle \alpha, \mu \rangle_h &&\text{for all } \alpha \in S_h', \mu \in S_i, \\
\langle \widetilde{Q}_H \alpha, \mu \rangle_H &= \langle \alpha, I_h \mu \rangle_h, &&\text{for all } \alpha \in S_h', \mu \in S_H^0.
\end{aligned}
\tag{3.5}
$$

The additive Schwarz preconditioner $B_h : S_h' \mapsto S_h$ is given by

$$B_h = \sum_{i=1}^N A_i^{-1} Q_i + I_h A_H^{-1} \widetilde{Q}_H, \tag{3.6}$$

where I_h is as defined by (3.3) and (3.4). A functional $g \in S_h'$ is completely represented by its action on a nodal basis of S_h, say $\{\eta_j\}$. Indeed, in computations, g is represented by a vector whose components are $\langle g, \eta_i \rangle_h$ (just as the right hand side of stiffness matrix equation (2.1) represents the functional in the right hand side of (1.3)). In iterative solution of (2.4), say by the preconditioned conjugate gradient method, one is required to compute $B_h g$, given the vector with components $\langle g, \eta_i \rangle_h$. From (3.6), it is clear that to compute $B_h g$, we need to solve subdomain problems as well as a coarse grid problem, i.e., we need to solve

for $v_i \in S_i$, and $v_H \in S_H^0$, given by

$$a_h(v_i, \mu) = \langle g, \mu \rangle_h \quad \text{for all } \mu \in S_i, \quad \text{and}$$
$$a_H(v_H, \mu) = \langle g, I_h\mu \rangle_h \quad \text{for all } \mu \in S_H.$$

Then $B_h g = \sum_{i=1}^N v_i + I_h v_H$. The expense of such a computation is justifed whenever the subdomain and coarse problems are small enough to permit their fast solution. Note that implementing the action of operators Q_i and $\tilde{Q}_H$ in (3.6) do not require Gramm matrix inversions.

The following theorem proves that B_h is a uniform preconditioner. The next section is devoted to a proof of this result.

Theorem 3.1. *The spectral condition number of the preconditioned operator $B_h A_h$ is bounded independently of h and H.*

4. Analysis of the preconditioner

In this section we prove Theorem 3.1. We use tools from previous analysis of Schwarz algorithms (cf. [16, 17, 24]), but a number of changes are necessitated due to our nontrivial intergrid transfer operator and mesh dependent bilinear form. The following four lemmas allow us to prove Theorem 3.1.

Lemma 4.1. *$B_h^{-1} : S_h \mapsto S_h'$ exists and for all $\lambda \in S_h$,*

$$\langle B_h^{-1}\lambda, \lambda \rangle_h = \min_{\lambda_i, \lambda_H} \left(\sum_{i=1}^N a_h(\lambda_i, \lambda_i) + a_H(\lambda_H, \lambda_H) \right),$$

where the minimum is taken over all decompositions of λ of the form

$$\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_N + I_h\lambda_H,$$

with $\lambda_i \in S_i$ and $\lambda_H \in S_H^0$.

Lemma 4.2. *For any $\lambda \in S_h$, there exists a decomposition $\lambda = \sum_{i=1}^N \lambda_i$ with $\lambda_i \in S_i$ such that*

$$C \sum_{i=1}^N a_h(\lambda_i, \lambda_i) \leqslant H^{-2}\|\lambda\|_h^2 + a_h(\lambda, \lambda).$$

Lemma 4.3. *For all $\lambda \in S_H^0$,*

$$a_h(I_h\lambda, I_h\lambda) \leqslant Ca_H(\lambda, \lambda).$$

Lemma 4.4. *For any $\lambda \in S_h$ there exists a $\lambda_H \in S_H^0$ such that*

$$\|\lambda - I_h\lambda_H\|_h \leqslant CHa_h(\lambda, \lambda)^{1/2}, \quad and \tag{4.1}$$
$$a_H(\lambda_H, \lambda_H) \leqslant Ca_h(\lambda, \lambda). \tag{4.2}$$

Before we prove these lemmas, let us show how the theorem follows from them.

Proof. (*Proof of Theorem 3.1.*) First let us prove that the smallest eigenvalue of $B_h A_h$ is uniformly bounded away from zero. Let $\lambda \in S_h$, and λ_H be as given by Lemma 4.4 applied to λ. Furthermore, let η_i be as given by Lemma 4.2 applied to $\lambda - I_h \lambda_H$. Then

$$\lambda - I_h \lambda_H = \sum_{i=1}^{N} \eta_i,$$

and by the estimates of Lemmas 4.3 and 4.4,

$$\sum_{i=1}^{N} \langle A_i \eta_i, \eta_i \rangle_h \leqslant C \left(H^{-2} \|\lambda - I_h \lambda_H\|_h^2 + a_h(\lambda - I_h \lambda_H, \lambda - I_h \lambda_H) \right)$$

$$\leqslant C a_h(\lambda, \lambda).$$

Thus we have found a decomposition $\lambda = \sum_{i=1}^{N} \eta_i + I_h \lambda_H$, with $\eta_i \in S_i$, and $\lambda_H \in S_H^0$ such that

$$\sum_{i=1}^{N} \langle A_i \eta_i, \eta_i \rangle_h + a_H(\lambda_H, \lambda_H) \leqslant C a_h(\lambda, \lambda).$$

Consequently by Lemma 4.1, $\langle B_h^{-1} \lambda, \lambda \rangle_h \leqslant C \langle A_h \lambda, \lambda \rangle_h$, and the assertion on the minimum eigenvalue of $B_h A_h$ follows.

It now remains to prove that the spectrum of $B_h A_h$ is bounded independently of h and H. We prove this by establishing that

$$a_h(B_h A_h \lambda, \lambda) \leqslant C a_h(\lambda, \lambda) \quad \text{for all } \lambda \in S_h. \tag{4.3}$$

Introducing operators $P_i : S_h \mapsto S_i$ and $\widetilde{P}_H : S_h \mapsto S_H$ defined by

$$a_h(P_i \lambda, \mu) = a_h(\lambda, \mu) \quad \text{for all } \lambda \in S_h, \mu \in S_i \quad \text{and}$$
$$a_H(\widetilde{P}_H \lambda, \mu) = a_h(\lambda, I_h \mu), \quad \text{for all } \lambda \in S_h, \mu \in S_H,$$

and observing that $Q_i A_h = A_i P_i$ and $\widetilde{Q}_H A_h = A_H P_H$ we get that

$$B_h A_h = \sum_{i=1}^{N} P_i + I_h \widetilde{P}_H. \tag{4.4}$$

The required upper bound involving $B_h A_h$ will follow if we show that $\widetilde{P}_H$ and P_i are bounded in $a_H(\cdot, \cdot)^{1/2}$ and $a_h(\cdot, \cdot)^{1/2}$ norms respectively. Since

$$a_H(\widetilde{P}_H \lambda, \widetilde{P}_H \lambda) = a_h(\lambda, I_h \widetilde{P}_H \lambda)$$
$$\leqslant a_h(\lambda, \lambda)^{1/2} a_h(I_h \widetilde{P}_H \lambda, I_h \widetilde{P}_H \lambda)^{1/2},$$

by Lemma 4.3 it follows that

$$a_H(\widetilde{P}_H \lambda, \widetilde{P}_H \lambda) \leqslant C a_h(\lambda, \lambda). \tag{4.5}$$

A similar bound also holds for P_i. Indeed,

$$a_h(P_i\lambda, P_i\lambda) = \int_{\widetilde{\Omega}_i} c\,\boldsymbol{\mathcal{Q}}(P_i\lambda) \cdot \boldsymbol{\mathcal{Q}}\lambda \, dx$$

$$\leqslant Ca_h(P_i\lambda, P_i\lambda)^{1/2} \left(\int_{\widetilde{\Omega}_i} c\,|\boldsymbol{\mathcal{Q}}\lambda|^2 \, dx \right)^{1/2},$$

so by Theorem 2.2 we get

$$a_h(P_i\lambda, P_i\lambda) \leqslant C \|\!|\lambda|\!\|_{h,\widetilde{\Omega}_i}^2. \tag{4.6}$$

Now, by usual arguments involving Assumption (3.2), the estimates (4.5) and (4.6) together with the identity (4.4) implies (4.3). □

In the remainder of this section, we prove Lemmas 4.2, 4.3, and 4.4. Proof of Lemma 4.1 involves only minor modifications of well known arguments and we omit it (cf. [5, 16, 21, 22]). The proof of Lemma 4.2 we now give is based on the standard Schwarz analyses [15, 16], so we will be brief.

Proof. (*Proof of Lemma 4.2.*) There exists a partition of unity $\{\theta_i(x)\}$, $0 \leqslant \theta_i(x) \leqslant 1$, such that the support of θ_i is contained in $\widetilde{\Omega}_i$, $\theta_i(x)$ is infinitely differentiable,

$$\sum_{i=1}^{N} \theta_i(x) = 1, \quad \text{and} \quad \|\boldsymbol{\nabla}\theta_i\|_{L^\infty(\widetilde{\Omega}_i)} \leqslant CH.$$

It is well known that the last inequality bounding the $L^\infty(\widetilde{\Omega}_i)$-norm of $\boldsymbol{\nabla}\theta_i$ holds due to Assumption (3.1). Furthermore, there exists a nodal interpolant of S_h, which we denote by Π_h, satisfying $\Pi_h\lambda|_{\partial\tau} = \lambda|_{\partial\tau}$ for all $\lambda \in S_h$, and $\|\Pi_h u\|_{h,\tau} \leqslant C\|u\|_{h,\tau}$ for all continuous functions u on $\partial\tau$. Define $\lambda_i \in S_i$ by

$$\lambda_i = \Pi_h(\theta_i\lambda).$$

Clearly, $\sum_{i=1}^{N} \lambda_i = \lambda$. We will now show that this decomposition satisfies the estimate of the lemma.

Let θ_i^τ denote the average of θ_i on any $\tau \in \mathcal{T}_h$, i.e., $\theta_i^\tau = |\tau|^{-1}\int_\tau \theta_i \, dx$. Then, using (2.10), the identity $\Pi_h(\theta_\tau^i\lambda)|_{\partial\tau} = \lambda|_{\partial\tau}$, and the approximation properties of averages,

$$
\begin{aligned}
\|\!|\Pi_h(\theta_i\lambda)|\!\|_{h,\tau} &\leqslant \|\!|\Pi_h(\theta_i - \theta_i^\tau)\lambda|\!\|_{h,\tau} + \|\!|\Pi_h(\theta_\tau^i\lambda)|\!\|_{h,\tau} \\
&\leqslant Ch^{-1}\|(\theta_i - \theta_i^\tau)\lambda\|_{h,\tau} + \|\!|\lambda|\!\|_{h,\tau} \\
&\leqslant Ch^{-1}\|\theta_i - \theta_i^\tau\|_{L^\infty(\tau)}\|\lambda\|_{h,\tau} + \|\!|\lambda|\!\|_{h,\tau} \\
&\leqslant Ch^{-1}(h\|\boldsymbol{\nabla}\theta_i\|_{L^\infty(\tau)})\|\lambda\|_{h,\tau} + \|\!|\lambda|\!\|_{h,\tau} \\
&\leqslant CH^{-1}\|\lambda\|_{h,\tau} + \|\!|\lambda|\!\|_{h,\tau}
\end{aligned}
$$

Squaring and summing over triangles $\tau \in \mathcal{T}_h$, we obtain the required estimate. □

To prove Lemmas 4.3 and 4.4, we first establish separately the following estimates for I_h^V.

Lemma 4.5. *The operator* $I_h^V : V_H^0 \mapsto V_h$ *satisfies*

$$|I_h^V w|_{H^1(\mathcal{T}_h)} \leqslant C|w|_{H^1(\mathcal{T}_H)}, \quad \text{and} \tag{4.7}$$

$$\|I_h^V w - w\|_{L^2(\Omega)} \leqslant CH|w|_{H^1(\mathcal{T}_H)}, \quad \text{for all } w \in V_H^0. \tag{4.8}$$

Proof. We first show the second estimate. Let z_i be an interior vertex of $\mathcal{T}_H$, and τ be a triangle connected to z_i. (In general, $\tau \neq \tau(z_i)$.) Then,

$$(I_h^V w - w)_{|\tau}(z_i) \equiv \lim_{\substack{z \to z_i \\ z \in \tau(z_i)}} I_h^V w(z) - w(z)$$

can be expressed as a telescoping sum of jumps of w across a few of the edges connected to z_i evaluated at z_i. Let $[w]_e(y)$ be the function defined for all $y \in e$ as the jump of w across e. (Its sign will not matter in the ensuing arguments.) Let $[w]_e(z_i) = \lim_{y \to z_i} [w]_e(y)$. Then,

$$\left| (I_h^V w - w)_{|\tau}(z_i) \right| \leqslant C \sum_{e \in \mathcal{E}(i)} \left| [w]_e(z_i) \right|^2,$$

where $\mathcal{E}(i)$ is the set of all edges connected to z_i. Consequently,

$$\|I_h^V w - w\|_{L^2(\Omega)}^2 \leqslant CH^2 \sum_{e \in \mathcal{E}_h} \left| [w]_e(z_i) \right|^2,$$

as cardinalities of $\mathcal{E}(i)$ are bounded independently of H.

Now, since $w \in V_H^0$, the jump $[w]_e(y)$ has zero average for all edges $e \in \mathcal{E}_h$. If $\tau_+(e)$ and $\tau_-(e)$ are the two triangles that share the edge e, then

$$\| [w]_e \|_{L^2(e)}^2 \leqslant CH \left(|\nabla w|_{L^2(\tau_+(e))}^2 + |\nabla w|_{L^2(\tau_-(e))}^2 \right).$$

Therefore,

$$\|I_h^V w - w\|_{L^2(\Omega)}^2 \leqslant CH \sum_{e \in \mathcal{E}_h} \| [w]_e \|_{L^2(e)}^2 \leqslant CH^2 |w|_{H^1(\mathcal{T}_H)}^2,$$

thus proving (4.8).

Estimate (4.7) follows from (4.8) as we now show:

$$\begin{aligned}
|I_h^V w|_{H^1(\mathcal{T}_h)}^2 = |I_h^V w|_{H^1(\mathcal{T}_H)}^2 \\
\leqslant \sum_{\tau \in \mathcal{T}_H} 2 \left(\|\nabla(I_h^V w - w)\|_{L^2(\tau)}^2 + \|\nabla w\|_{L^2(\tau)}^2 \right) \\
\leqslant CH^{-2} \|I_h^V w - w\|_{L^2(\Omega)}^2 + 2|w|_{H^1(\mathcal{T}_H)}^2 \\
\leqslant C|w|_{H^1(\mathcal{T}_H)}^2,
\end{aligned}$$

by using a standard inverse inequality and (4.8).   □

Proof. (*Proof of Lemma 4.3.*) For any $\eta \in S_H^0$, using Lemmas 3.1 and 4.5, we have

$$\begin{aligned}
\|I_h \eta\|_h^2 = \|X_h^{-1} I_h^V X_H \eta\|_h^2 \\
\leqslant C|I_h^V X_H \eta|_{H^1(\mathcal{T}_h)}^2 \leqslant C|X_H \eta|_{H^1(\mathcal{T}_H)}^2 \\
\leqslant C\|\eta\|_H^2.
\end{aligned}$$

□

To prove Lemma 4.4, we need the following additional result.

Lemma 4.6. *The L^2-orthogonal projection Q_H^V into V_H^0 satisfies*

$$|Q_H^V w|_{H^1(\mathcal{T}_H)} \leqslant C|w|_{H^1(\mathcal{T}_h)} \quad and \tag{4.9}$$

$$\|w - Q_H^V w\|_{L^2(\Omega)} \leqslant CH|w|_{H^1(\mathcal{T}_h)} \quad for \ all \ w \in V_h^0. \tag{4.10}$$

Proof. Let $W_h \equiv \{ v : v$ is continuous on Ω, v is zero on $\partial\Omega$, and v is linear on every $\tau \in \mathcal{T}_h \}$. By a straightforward modification of the proof of Lemma 4.5 we can prove that there is a $\bar{w} \in W_h$ such that

$$\|w - \bar{w}\|_{L^2(\Omega)} \leqslant Ch|w|_{H^1(\mathcal{T}_h)}$$
$$\|\boldsymbol{\nabla} w\|_{0,\Omega} \leqslant C|w|_{H^1(\mathcal{T}_h)}$$

It is easy to see, e.g., by using the well known properties of the L^2-orthogonal projection into W_H, that there exists a $\bar{w}_H \in W_H$ such that

$$\|w - \bar{w}_H\|_{L^2(\Omega)} \leqslant CH\|\boldsymbol{\nabla}\bar{w}\|_{L^2(\Omega)}$$
$$\|\boldsymbol{\nabla}\bar{w}_H\|_{L^2(\Omega)} \leqslant C\|\boldsymbol{\nabla}\bar{w}\|_{L^2(\Omega)}.$$

Therefore,

$$\begin{aligned}
\|w - Q_H^V w\|_{L^2(\Omega)} &\leqslant \|w - \bar{w}_H\|_{L^2(\Omega)} \\
&\leqslant \|w - \bar{w}\|_{L^2(\Omega)} + \|\bar{w} - \bar{w}_H\|_{L^2(\Omega)} \\
&\leqslant CH|w|_{H^1(\mathcal{T}_h)}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
|Q_H^V w|_{H^1(\mathcal{T}_H)} &\leqslant |Q_H^V w - \bar{w}_H|_{H^1(\mathcal{T}_H)} + |\bar{w}_H|_{H^1(\mathcal{T}_H)} \\
&\leqslant CH^{-1}\|Q_H^V w - \bar{w}_H\|_{L^2(\Omega)} + C|w|_{H^1(\mathcal{T}_h)} \\
&\leqslant CH^{-1}\left(\|Q_H^V w - w\|_{L^2(\Omega)} + \|w - \bar{w}_H\|_{L^2(\Omega)}\right) + C|w|_{H^1(\mathcal{T}_h)} \\
&\leqslant C|w|_{H^1(\mathcal{T}_h)}.
\end{aligned}$$

$\square$

Proof. (*Proof of Lemma 4.4.*) For any $\lambda \in S_h$ we set $\lambda_H = X_H^{-1}Q_H^V X_h Q_h^{S,0}\lambda$ where $Q_h^{S,0}$ is the L^2-orthogonal projection into S_h^0. We will now show that the λ_H so defined satisfies both the estimates of the lemma, namely (4.1) and (4.2). First, note that the $Q_h^{S,0}$ satisfies

$$\|Q_h^{S,0}\lambda\|_h \leqslant C\|\lambda\|_h, \quad and \tag{4.11}$$

$$\|\lambda - Q_h^{S,0}\lambda\|_h \leqslant Ch\|\lambda\|_h, \tag{4.12}$$

for all $\lambda \in S_h$. Both these estimates follow from straightforward scaling arguments and the observation that whenever λ is a constant along the perimeter of a triangle $\tau \in \mathcal{T}_h$, $Q_h^{S,0}\lambda$ coincides with λ on $\partial\tau$.

To prove (4.2), we use Lemmas 4.6 and 3.1, and (4.11):

$$\begin{aligned}
\|\lambda_H\|_H &\leqslant C|Q_H^V X_h Q_h^{S,0}\lambda|_{H^1(\mathcal{T}_H)} \\
&\leqslant C|X_h Q_h^{S,0}\lambda|_{H^1(\mathcal{T}_h)} \\
&\leqslant C\|Q_h^{S,0}\lambda\|_h \leqslant C\|\lambda\|_h.
\end{aligned}$$

Estimate (4.2) now follows from Theorem 2.2.

To prove (4.1), we start by using (4.12)

$$\|\lambda - I_h \lambda_H\|_h \leqslant \|\lambda - Q_h^{S,0}\lambda\|_h + \|Q_h^{S,0}\lambda - I_h \lambda_H\|_h$$
$$\leqslant Ch\|\lambda\|_h + \|Q_h^{S,0}\lambda - I_h \lambda_H\|_h$$

We now estimate the last term. By Lemma 3.1,

$$\|Q_h^{S,0}\lambda - I_h \lambda_H\|_h = \|X_h^{-1}(X_h Q_h^{S,0}\lambda - I_h^V X_H \lambda_H)\|_h$$
$$\leqslant C\|X_h Q_h^{S,0}\lambda - I_h^V X_H \lambda_H\|_h.$$

Moreover, applying Lemmas 4.6, 4.5, and 3.1,

$$\|X_h Q_h^{S,0}\lambda - I_h^V X_H \lambda_H\|_h \leqslant \|X_h Q_h^{S,0}\lambda - X_H \lambda_H\|_h + \|X_H \lambda_H - I_h^V X_H \lambda_H\|_h$$
$$= \|(I - Q_H^V)X_h Q_h^{S,0}\lambda\|_h + \|(I - I_h^V)X_H \lambda_H\|_h$$
$$\leqslant CH|X_h Q_h^{S,0}\lambda|_{H^1(\mathcal{T}_h)} + CH|X_H \lambda_H|_{H^1(\mathcal{T}_H)}$$
$$\leqslant CH\|Q_h^{S,0}\lambda\|_h + CH\|\lambda_H\|_H$$

An application of (4.11) and (4.2) now completes the proof. $\square$

5. Concluding remarks

We now briefly mention a few corollaries of our analysis. The I_h^V we introduced in this paper can be used as an intergrid transfer operator in an additive Schwarz algorithm to define a preconditioner for the P1-nonconforming method. It can be proved, either by the general strategy here, or by verifying conditions stated in [7], that the resulting preconditioner is uniform with respect to fine and coarse mesh sizes. The critical estimates involved are those given by Lemma 4.5.

We have shown how the intergrid transfer operator I_h^V between the P1-nonconforming spaces can be combined with the isomorphism X_h to yield intergrid transfer operators suitable for the hybridized mixed method. The analysis continues to hold if our I_h^V is substituted with some other intergrid transfer operators for the P1-nonconforming method developed elsewhere. More precisely, abstracting the properties of I_h^V used in our analysis, we have the following theorem:

Theorem 5.1. *If $I_h = X_h^{-1}\widetilde{I}_h^V X_H$ for some $\widetilde{I}_h^V : V_H^0 \mapsto V_h^0$ satisfying*

$$|\widetilde{I}_h^V w|_{H^1(\mathcal{T}_h)} \leqslant C|w|_{H^1(\mathcal{T}_H)}, \quad \text{and} \tag{5.1}$$

$$\|\widetilde{I}_h^V w - w\|_{L^2(\Omega)} \leqslant CH|w|_{H^1(\mathcal{T}_H)}, \quad \text{for all } w \in V_H^0, \tag{5.2}$$

and if $\widetilde{Q}_H$ is as defined by (3.5), then the operator B_h defined by (3.6) is a uniform preconditioner for A_h.

To consider a few applications of this theorem, let $w \in V_H^0$, and let z_i be a vertex of $\mathcal{T}_H$. In general, on different triangles $\tau \in \mathcal{T}_H$ connected to z_i, the limit

$$\lim_{\substack{z \to z_i \\ z \in \tau}} w(z)$$

will differ. Let us denote by w_i the average of these limiting values.

Example 5.1. The following $\widetilde{I}_h^V$ was defined in [7]: $\widetilde{I}_h^V w$ is the unique function in V_h^0 whose values at the midpoints of $e \in \mathcal{E}_h$ coincide with those of the function $\widetilde{w}$ that is continuous on Ω, zero on $\partial\Omega$, quadratic on every $\tau \in \mathcal{T}_H$, equals w at midpoints of $e \in \mathcal{E}_H$, and equals w_i at vertices z_i of $\mathcal{T}_H$. It is proved in [7] that it satisfies (5.1) and (5.2).

Example 5.2. The following intergrid transfer operator was defined in [3] for use in a multilevel algorithm: $\widetilde{I}_h^V w$ is the unique function in V_h^0 satisfying

$$\int_\Omega (\widetilde{I}_h^V w) v_h = \int_\Omega w v_h \, dx \quad \text{for all } v_h \in V_h^0.$$

It follows from Lemma 4.5 that $\widetilde{I}_h^V$ satisfies (5.1):

$$\|\widetilde{I}_h^V w - w\|_{L^2(\Omega)}^2 = \int_\Omega (\widetilde{I}_h^V w - w)(I_h^V w - w) \, dx$$
$$\leqslant CH \|\widetilde{I}_h^V w - w\|_{L^2(\Omega)} |w|_{H^1(\mathcal{T}_H)}.$$

Proof of (5.2) is similar to that of (4.7). Note that implementation of this intergrid transfer operator requires Gramm matrix inversions.

Example 5.3. The I_h^V we have introduced (see (3.4)) is computationally less expensive than the intergrid transfers of the previous two examples. However, it does not use information from all available coarse grid degrees of freedom. Therefore, we introduce the following modified operator: $\widetilde{I}_h^V w$ is the unique function in W_H which equals w_i at all interior vertices z_i of $\mathcal{T}_H$. By arguments essentially similar to those in the proof of Lemma 4.5, we can prove that (5.1) and (5.2) holds for this $\widetilde{I}_h^V$.

Unfortunately, it is difficult to assert using the analysis of this paper, that any one of the intergrid transfer operators above is better than another. Such a comparison is probably best done computationally.

It is possible to consider higher order coarse spaces instead of S_H^0. Obviously, in such cases, the isomorphism with lowest order nonconforming space that we used here will not suffice. An analysis in the higher order case of a multilevel algorithm can be found in a sequel.

Acknowledgment

References

[1] D. N. ARNOLD, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[2] D. N. ARNOLD AND F. BREZZI, *Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 7–32.

[3] D. BRAESS AND R. VERFÜRTH, *Multigrid methods for nonconforming finite element methods*, SIAM J. Numer. Anal., 27 (1990), pp. 979–986.

[4] 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.

[5] J. H. Bramble and X. Zhang, *The analysis of multigrid methods*, in Handbook of numerical analysis, Vol. VII, North-Holland, Amsterdam, 2000, pp. 173–415.

[6] S. C. Brenner, *A multigrid algorithm for the lowest-order Raviart-Thomas mixed triangular finite element method*, SIAM J. Numer. Anal., 29 (1992), pp. 647–678.

[7] S. C. Brenner, *Two-level additive Schwarz preconditioners for nonconforming finite element methods*, Math. Comp., 65 (1996), pp. 897–921.

[8] S. C. Brenner, *Poincaré-Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal., (To appear).

[9] 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.

[10] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, no. 15 in Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.

[11] Z. Chen, *The analysis of intergrid transfer operators and multigrid methods for nonconforming finite elements*, Electron. Trans. Numer. Anal., 6 (1997), pp. 78–96 (electronic). Special issue on multilevel methods (Copper Mountain, CO, 1997).

[12] Z. Chen, R. E. Ewing, R. D. Lazarov, S. Maliassov, and Y. A. Kuznetsov, *Multilevel preconditioners for mixed methods for second order elliptic problems*, Numer. Linear Algebra Appl., 3 (1996), pp. 427–453.

[13] B. Cockburn and J. Gopalakrishnan, *A characterization of hybridized mixed methods for the Dirichlet problem*, Submitted to SIAM J. Numer. Anal., (2002). Available as IMA preprint 1897 on the web at `www.ima.umn.edu/preprints/`.

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

[15] M. Dryja, B. F. Smith, and O. B. Widlund, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, SIAM J. Numer. Anal, 31 (1994), pp. 1662–1694.

[16] M. Dryja and O. B. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*, Tech. Rep. 339, Courant Institute of Mathematical Sciences, New York, 1987.

[17] ——, *Some domain decompostion algorithms for elliptic problems*, in Iterative Methods for Large Linear Systems, Academic Press, San Diego, 1989, pp. 273–291. Proceedings of a conference held at Austin, Texas, in October 1988.

[18] J. Gopalakrishnan and G. Kanschat, *A multilevel discontinuous Galerkin method*, Numer. Math., (2002). To appear.

[19] Y. A. Kuznetsov and S. Y. Maliassov, *Substructuring preconditioners for nonconforming finite element approximations of second-order elliptic problems with anisotropy*, Russian J. Numer. Anal. Math. Modelling, 10 (1995), pp. 511–533.

[20] R. D. Lazarov, J. E. Pasciak, and P. S. Vassilevski, *Iterative solution of a coupled mixed and standard Galerkin discretization method for elliptic problems*, Numer. Linear Algebra Appl., 8 (2001), pp. 13–31.

[21] P. L. Lions, *On the Schwarz alternating method*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., SIAM, Philadelphia, 1988, pp. 1–42. Symposium held at Ecole Nationale des Ponts et Chaussées, Paris, in January 1987.

[22] P. Oswald, *Intergrid transfer operators and multilevel preconditioners for nonconforming discretizations*, Appl. Numer. Math., 23 (1997), pp. 139–158. Multilevel methods (Oberwolfach, 1995).

[23] P.-A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), Springer, Berlin, 1977, pp. 292–315. Lecture Notes in Math., Vol. 606.

[24] T. RUSTEN, P. S. VASSILEVSKI, AND R. WINTHER, *Interior penalty preconditioners for mixed finite element approximations of elliptic problems*, Math. Comp., 65 (1996), pp. 447–466.

[25] T. RUSTEN AND R. WINTHER, *Substructure preconditioners for elliptic saddle point problems*, Math. Comp., 60 (1993), pp. 23–48.