NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2009; **Volume N**:1–24 *Prepared using* nlaauth.cls *[Version: 2002/09/18 v1.02]*

A Convergent Multigrid Cycle for the Hybridized Mixed Method

Jayadeep Gopalakrishnan[1]*, Shuguang Tan[1]

[1] *University of Florida, Department of Mathematics, Gainesville, FL 32611–8105*

SUMMARY

We consider the application of a variable V-cycle multigrid algorithm for the hybridized mixed method for second order elliptic boundary value problems. Our algorithm differs from previous works on multigrid for the mixed method in that it is targeted at efficiently solving the matrix system for the Lagrange multiplier of the method. Since the mixed method is best implemented by first solving for the Lagrange multiplier and recovering the remaining unknowns locally, our algorithm is more useful in practice. The critical ingredient in the algorithm is a suitable intergrid transfer operator. We design such an operator and prove mesh independent convergence of the variable V-cycle algorithm. Numerical experiments indicating the asymptotically optimal performance of our algorithm, as well as the failure of certain seemingly plausible intergrid transfer operators, are presented. Copyright © 2009 John Wiley & Sons, Ltd.

KEY WORDS: multigrid, hybrid, mixed method, Raviart-Thomas, BDM, HDG, nonnested

1. Introduction

The mixed method is an important finite element method for numerically solving partial differential equations on complicated domains. This paper presents a multigrid method for solving the linear systems arising from the so-called hybridized mixed method [1, 14]. Like other finite element methods, the hybridized mixed method yields matrix systems with condition number that grows as mesh size decreases. Hence it is necessary to use preconditioned iterative solvers or fast linear solvers like multigrid algorithms to obtain the solution efficiently. Classical iterative methods such as the Gauss-Seidel iteration, reduce high frequency components of the iterative error quickly, but one needs techniques like the multigrid V-cycle iteration to reduce all components of the error efficiently. It is well known that multigrid techniques can give solvers of asymptotically optimal complexity for many applications. We shall show that this is the case for the hybridized mixed method as well, provided one chooses the ingredients appropriately in a V-cycle algorithm. Although the subject of efficient solvers for the mixed method has been actively pursued by many authors, as we shall describe shortly, our algorithm addresses

Contract/grant sponsor: This work was supported in part by the National Science Foundation under grants DMS-0713833 and SCREMS-0619080.

Received September 3, 2008
Revised January 1, 2009
Accepted ¡Date¿

a practical need better than the existing algorithms. Furthermore, to put this work in the perspective of recent developments in the design of new hybridized DG (HDG) methods [16], our algorithm can be thought of as providing an efficient solver for the simplest HDG method, namely the hybridized mixed method. Thus this contribution represents a first step towards the design of efficient solvers for all HDG methods.

We consider the hybridized mixed method for the Dirichlet problem

$$\begin{aligned} -\Delta u &= f \text{ on } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{1}$$

where Ω is a polygonal domain and $f \in L^2(\Omega)$. We mesh Ω by a finite element triangulation and derive the method by reformulating the above system, after introducing the flux variable $\boldsymbol{q}$, as

$$\begin{aligned} \boldsymbol{q} &= -\nabla u && \text{in } \Omega, \\ \operatorname{div}\boldsymbol{q} &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned}$$

multiplying by test functions and integrating by parts on each mesh element. The resulting system (which we shall describe precisely in Section 2) has three unknown variables, an approximation $\boldsymbol{q}_h$ for the exact flux $\boldsymbol{q}$, an approximation u_h for the primal variable u, and a Lagrange multiplier λ_h approximating the traces of the solution u on the element boundaries. The non-hybridized form of the mixed method, in contrast, has only $\boldsymbol{q}_h$ and u_h as unknowns.

While it may appear that the hybridized formulation has too many unknowns ($\boldsymbol{q}_h$, u_h, and λ_h, instead of just $\boldsymbol{q}_h$ and u_h), it has long been known [1, 14] that the hybridized formulation is better. Indeed, in the hybridized formulation, one can eliminate the unknowns $\boldsymbol{q}_h$ and u_h and obtain a global system solely involving λ_h. The size of this reduced system is smaller than the size of the original mixed method in the higher order case. Furthermore, this system is symmetric and positive definite, so excellent iterative techniques such as the Conjugate Gradients can be used. It is important to also note that once λ_h is computed, the other components of the solution, namely $\boldsymbol{q}_h$ and u_h can be computed element by element through local operations. Finally, it is also well known [1] that one can postprocess the hybrid solution to obtain another approximation with enhanced accuracy. Clearly, these are compelling reasons to prefer the hybridized formulation.

The above mentioned global system that involves solely the Lagrange multiplier λ_h is the system we must solve efficiently. Since the recovery of the remaining unknowns only involves local operations, the solution cost is typically dominated by the cost of solving the global system for λ_h. It is for this system that we need efficient solvers, like multigrid techniques. However, there are some difficulties in adapting a standard multigrid algorithm to this system, arising due to two non-standard features of hybridized methods: (1) The space in which λ_h lies consist of functions defined on the *edges* of the mesh. In contrast, typical finite element methods uses functions defined on the elements of a mesh. (2) The approximate solution λ_h satisfies a variational formulation involving a *mesh dependent* bilinear form. In contrast many of the standard finite element approximations are characterized via a variational equation that is defined in a Sobolev space containing the finite element spaces. Because of these features, when adapting multigrid algorithms to hybridized mixed methods, we must design the components of the algorithm to work with *non-nested* multilevel finite element spaces, and *non-inherited*

bilinear forms. For analysis, we use the abstract theory of [7] which has often proved successful in analyzing multigrid adaptations to other problems with similar difficulties.

The two main ingredients of any multigrid algorithms are smoothers and intergrid transfer operators. For the problems we shall consider, smoothers do not pose new challenges. However, the design of proper intergrid transfer operators tailored to the hybridized schemes are critical for the success of the multigrid algorithm. These are operators used within the algorithm to move data from a coarser grid to a finer grid. Since hybridized methods use functions defined on mesh edges, we must design non-trivial intergrid transfer operators. As our numerical experiments show, some of the "obvious" choices do *not* result in efficient, or even convergent, multigrid algorithms. We have chosen our intergrid transfers and multilevel spaces such that at the coarser levels we use the lowest order conforming finite element spaces, so that the number of degrees of freedom of the coarse levels are minimized. Conforming coarser spaces have been used previously in other contexts [7, 17].

There are a number of previous works on multigrid algorithms for mixed and hybridized mixed methods for the Poisson equation [7, 9, 22, 25]. One of the first is in fact contained in the same paper [7] that has the abstract multigrid theory we shall use here. They analyzed a multigrid algorithm for a matrix system involving the primal variable u alone. However the evaluation of the operator acting on u on general meshes requires solving a global system. Although this system is well conditioned, we find this approach practically inconvenient. The work [22] explores an interesting idea for preconditioning the system for u using a spectrally equivalent discontinuous Galerkin like bilinear form, while [23] applies domain decomposition ideas to precondition a system obtained by eliminating $\boldsymbol{q}_h$, but involving both u_h and λ_h. Our work differs from these in that we design a solver for the system involving λ_h alone. This is the most practically interesting strategy, because the best implementation of the mixed method proceeds by assembling the global system for λ_h alone, and recovering the remaining solutions components $\boldsymbol{q}_h$ and u_h locally after solving for λ_h. Hence what is needed most in practice is an efficient solver for the matrix system for λ_h. The paper [9] of Brenner does present a multigrid algorithm for the system involving λ_h alone, but only for the lowest order case, exploiting its connection with the P_1-nonconforming method. Since such connections are not known for the Lagrange multiplier system for the general order case, this approach cannot be generalized to the arbitrary order case. The same is true for other strategies like those [12, 13] that uses the equivalence with the P_1-nonconforming method. In our approach there is no difference between the lowest order and the remaining cases. A Schwarz preconditioner for the arbitrary order system for λ_h can be found in [19]. In this paper, we consider the V-cycle multigrid algorithm instead.

A few remarks on our method of analysis are in order. There are a number of frameworks [7, 10, 18] for analyzing multigrid algorithms similar to ours. We use the abstract theory of [7], whereby the convergence proof is reduced to a few conditions which we are able to verify. Our techniques also give results for W-cycle type algorithms, but in our opinion, a result on the convergence of the variable V-cycle is perhaps the most interesting practically, hence we shall only state such a result. In the proof of convergence, we bring to bear certain recent results on the hybridized mixed method [14, 15]. In particular, an error estimate for the Lagrange multiplier proved in [15] plays a critical role. Indeed, previously known error estimates [1] for λ_h (in the arbitrary degree case) required at least H^3-regularity of the exact solution, which makes them useless in regularity based multigrid analyses. We instead present a full regularity based multigrid proof using the new error estimate of [15]. We also develop a series

of intermediate lemmas which we anticipate to be useful in the analysis of multigrid for other similar applications like HDG methods.

In the next section we shall precisely describe the hybridized mixed method, more specifically the hybridized Raviart-Thomas method. In Section 3, we present the multigrid algorithm. The succeeding section is devoted to the convergence analysis of this algorithm. We conclude in Section 5 by reporting the results of our numerical experiments.

2. The hybridized mixed method

We now present the well known [1, 14] hybridized mixed method for the numerical solution of the Dirichlet problem (1) and list a few results regarding the method which we shall use in later sections.

Reformulating (1) as a first order system of two equations, the first equation is the definition of flux, namely $\boldsymbol{q} = -\nabla u$, while the second equation is $\mathrm{div}\boldsymbol{q} = f$. Multiplying the first equation by a test function $\boldsymbol{r}$ and integrating by parts we obtain

$$\int_K \boldsymbol{q}\cdot\boldsymbol{r} \; - \int_K u\,\mathrm{div}\boldsymbol{r} \; + \int_{\partial K} u\,\boldsymbol{r}\cdot\boldsymbol{n} = 0 \tag{2}$$

for any triangle K contained in Ω. Above and elsewhere, we drop the measure in integrals, as it will be clear from the context. Moreover, we use $\boldsymbol{n}$ to denote the unit outward normal on any domain – the specific domain under consideration in each instance will be clear from the context, e.g., above it is K. A discrete version of (2) forms the first equation of the hybridized method. To describe it, we need to define appropriate finite element spaces. Let $\mathscr{T}_h$ be a finite element triangulation of Ω satisfying the usual geometrical conformity constraints. The subscript h refers to the maximum of the diameters of all the triangles K in the collection $\mathscr{T}_h$. We assume that $\mathscr{T}_h$ is quasiuniform for simplicity. Let $P_d(K)$ denote the set of polynomials of degree at most d on K. Define the following spaces of piecewise polynomials:

$$\begin{aligned} V_h &= \{\boldsymbol{v} : \quad \boldsymbol{v}|_K \in R_d(K), \quad \text{for all } \; K \in \mathscr{T}_h\},\\ T_h &= \{w : \quad w|_K \in P_d(K), \quad \text{for all } \; K \in \mathscr{T}_h\}, \end{aligned}$$

where $R_d(K) = P_d(K)\times P_d(K) + \boldsymbol{x}P_d(K)$, the well known Raviart-Thomas space [21]. Let $\mathscr{E}_h$ denote the set of all edges in $\mathscr{T}_h$, and $\mathscr{E}_h^i$ be the set of interior edges (not contained in $\partial\Omega$). We need one more discrete space, this time a space of functions defined on the union of edges

$$M_h = \{\mu : \quad \mu|_e \in P_d(e), \quad \text{for all } \; e \in \mathscr{E}_h^i\}.$$

We also identify the functions in M_h with their extension by zero to the union of all mesh edges in $\mathscr{E}_h$, i.e., if we refer to the value of a $\mu \in M_h$ on $\partial\Omega$, it is understood to be zero. Therefore, we can write identities like

$$\sum_{K\in\mathscr{T}_h} \int_{\partial K} \mu\,\boldsymbol{r}\cdot\boldsymbol{n} \;=\; \sum_{e\in\mathscr{E}_h^i} \int_e \mu\,[\![\boldsymbol{r}\cdot\boldsymbol{n}]\!], \qquad \text{for all } \mu\in M_h,\; \boldsymbol{r}\in V_h, \tag{3}$$

where $[\![\boldsymbol{r}\cdot\boldsymbol{n}]\!]$ is the jump of the normal component of $\boldsymbol{r}$, defined as follows: On any edge e shared by two mesh triangles K^+ and K^-, the jump on e is defined by $[\![\boldsymbol{r}\cdot\boldsymbol{n}]\!] = \boldsymbol{r}|_{K^+}\cdot\boldsymbol{n}^+ + \boldsymbol{r}|_{K^-}\cdot\boldsymbol{n}^-$ where $\boldsymbol{n}^\pm$ denotes the outward unit normal on $\partial K^\pm$ (see Figure 1).

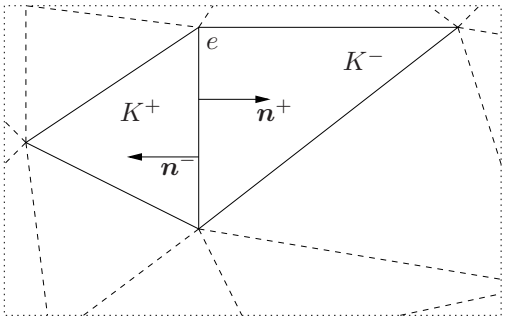

Figure 1. Illustration of notations

With these spaces and notations, we can now write down the hybridized Raviart-Thomas mixed method. Its approximation triple $\boldsymbol{q}_h \in V_h,\ u_h \in T_h,\ \lambda_h \in M_h$, is defined by requiring that, for all $(\boldsymbol{v}, w, \mu) \in V_h \times T_h \times M_h$,

$$\int_\Omega \boldsymbol{q}_h \cdot \boldsymbol{v} - \sum_{K\in\mathscr{T}_h} \int_K u_h \mathrm{div}\boldsymbol{v} + \sum_{e\in\mathscr{E}_h^i} \int_e \lambda_h \ [\![\boldsymbol{v}\cdot\boldsymbol{n}]\!] = 0, \tag{4a}$$

$$\sum_{K\in\mathscr{T}_h} \int_K w \mathrm{div}\boldsymbol{q}_h = \int_\Omega f\, w, \tag{4b}$$

$$\sum_{e\in\mathscr{E}_h^i} \int_e \mu \ [\![\boldsymbol{q}_h\cdot\boldsymbol{n}]\!] = 0. \tag{4c}$$

It is easy to show that this system has a unique solution [15]. Comparing the first equation with (2), keeping (3) in view, we expect λ_h to be an approximation to the trace of the exact solution u on $\mathscr{E}_h^i$. This is indeed the case as is quantified in the error estimates proved in [1, 15]. For our multigrid analysis later, the older error estimate of [1] is unsuitable, and the newer estimate of [15] plays a critical role.

It is well known that the (nonhybridized) Raviart-Thomas mixed method as originally presented in [21], involving just $\boldsymbol{q}_h$ and u_h as unknowns, is equivalent to the above hybridized mixed method, in that the flux and primal variable approximations ($\boldsymbol{q}_h$ and u_h) of both methods coincide. Hence (4) is often thought of as an alternate implementation technique for the mixed method. Although the hybridized method is usually presented as in (4), it is not advisable to implement it in that form. Indeed, the main advantage of hybridization is that all the unknowns, except λ_h, can be eliminated to get one global system for λ_h. Then $\boldsymbol{q}_h$ and u_h are locally recoverable from λ_h. We shall now describe this precisely in the next theorem, after establishing some necessary notations.

We introduce two local operators. The first mapping lifts functions on edges of the triangulation $\mathscr{T}_h$ to functions on Ω. Let $L^2(\mathscr{E}_h)$ denote the set of all square integrable functions on the union of all edges of $\mathscr{E}_h$. For any $\mu \in L^2(\mathscr{E}_h)$, the local liftings $\mathfrak{Q}\mu$ and $\mathfrak{U}\mu$ are functions

in V_h and T_h respectively, defined by requiring that

$$\begin{aligned} \int_\Omega \mathcal{Q}\mu \cdot \boldsymbol{v} - \sum_{K\in\mathscr{T}_h} \int_K \mathcal{U}\mu \,\mathrm{div}\boldsymbol{v} &= -\sum_{e\in\mathscr{E}_h} \int_e \mu\, [\![\boldsymbol{v}\cdot\boldsymbol{n}]\!], \\ \sum_{K\in\mathscr{T}_h} \int_K w\,\mathrm{div}\mathcal{Q}\mu &= 0, \end{aligned} \tag{5}$$

hold for all $(\boldsymbol{v}, w) \in V_h \times T_h$. Note that since the functions above have *no* continuity constraints across elements, the computation of $(\mathcal{Q}\mu, \mathcal{U}\mu)$ can be done element by element in a decoupled way. Such computations, being *local*, are inexpensive. The second mapping, associates to the function $f \in L^2(\Omega)$ the element $(\mathcal{Q}f, \mathcal{U}f) \in V_h \times T_h$ defined by

$$\begin{aligned} \int_\Omega \mathcal{Q}f \cdot \boldsymbol{v} - \sum_{K\in\mathscr{T}_h} \int_K \mathcal{U}f \mathrm{div}\boldsymbol{v} &= 0, \\ \sum_{K\in\mathscr{T}_h} \int_K w\,\mathrm{div}\mathcal{Q}f &= \int_\Omega f\, w, \end{aligned} \tag{6}$$

for all $(\boldsymbol{v}, w) \in V_h \times T_h$. These mappings are uniquely defined because of the surjectivity of the map $\mathrm{div} : V_h \mapsto T_h$. With these maps, we have the following theorem [14].

Theorem 2.1. *Let $(\boldsymbol{q}_h, u_h, \lambda_h)$ be the solution of the hybridized Raviart-Thomas method (4). Then*

$$\boldsymbol{q}_h = \mathcal{Q}\lambda_h + \mathcal{Q}f \qquad \textit{and} \qquad u_h = \mathcal{U}\lambda_h + \mathcal{U}f. \tag{7}$$

The Lagrange multiplier $\lambda_h \in M_h$ is the unique solution of

$$a(\lambda_h, \mu) = b(\mu) \quad \textit{ for all } \mu \in M_h, \tag{8}$$

where

$$a(\lambda_h, \mu) = \int_\Omega \mathcal{Q}\lambda_h \cdot \mathcal{Q}\mu$$

and

$$b(\mu) = \int_\Omega f\, \mathcal{U}\mu.$$

By virtue of this theorem, instead of dealing with (4) directly, we can solve (8) to find the Lagrange multiplier λ_h, and then use (7) to recover the $\boldsymbol{q}_h$ and u_h. This recovery is *local*, as the application of $\mathcal{Q}$ and $\mathcal{U}$ are local operations, and consequently of negligible cost, compared to the global inversion required to find λ_h.

Remark 2.1. The BDM method [11] can also be hybridized along similar lines and a result similar to Theorem 2.1 holds for it. In fact, in [14] it is proved that the variational equation for the λ_h from the BDM method and the RT method has the *same* bilinear form $a(\cdot,\cdot)$. Consequently, the multigrid algorithm for $a(\cdot,\cdot)$ that we give next, applies verbatim to the hybridized BDM method.

3. The multigrid algorithm

In this section, we will give our multigrid algorithm for solving the linear system arising from the hybridized mixed method. As we have already seen, all solution components can be recovered once we find λ_h in M_h satisfying

$$a(\lambda_h, \mu) = b(\mu), \qquad \text{for all } \mu \in M_h.$$

The stiffness matrix of this mesh dependent variational system has condition number that grows like $O(h^{-2})$ as proved in [19]. Therefore, we need iterative solution strategies that do not deteriorate in performance when condition number increases. In this section, we give such an iterative scheme using multigrid techniques. The algorithm fits into the abstract framework of [7, 8] as one of their abstract variable V-cycle algorithms. Their algorithm, with the abstract components particularized to our application is given below, following which we state our main result on the convergence of the algorithm.

Multigrid algorithms require a multilevel hierarchy of meshes and spaces, which we now describe. We assume that the mesh $\mathscr{T}_h$ in which the solution is sought, is obtained by successive refinements of a coarse mesh $\mathscr{T}_1$. At a refinement level $k = 2, 3, \ldots, J$, the mesh $\mathscr{T}_k$ is obtained from $\mathscr{T}_{k-1}$ by connecting the midpoints of all edges of $\mathscr{T}_{k-1}$. Let $\mathscr{E}_k$ denote the set of all interior edges of $\mathscr{T}_k$. By an abuse of notation, the domain formed by the union of all mesh edges in $\mathscr{E}_k$, is also denoted by the same $\mathscr{E}_k$. Let h_k denote the mesh size of $\mathscr{T}_k$, so in particular, $h \equiv h_J$. Let us now define the *multilevel spaces*. Define M_k by

$$M_k = \{v : \Omega \mapsto \mathbb{R} \big| \ v \text{ is continuous, } v|_{\partial\Omega} = 0, \ v|_K \in P_1(K), \ \forall \text{ triangles } K \in \mathscr{T}_{k+1}\},$$

for $k = 0, 1, \ldots, J-1$ and define

$$M_J = \{\mu : \mathscr{E}_J \to \mathbb{R} \big| \quad \mu|_e \in P_d(e), \forall e \in \mathscr{E}_J\}$$

where d is a nonnegative integer. Note that we have $J+1$ spaces here and the final space is where the Lagrange multiplier solution λ_h lies, i.e.,

$$M_J = M_h.$$

Note also that although $M_0 \subseteq M_1 \subseteq \cdots \subseteq M_{J-1}$, the last space is not nested, i.e., $M_{J-1} \not\subseteq M_J$. Hence, we must develop multigrid algorithms in a non-nested space setting.

Each M_k is endowed with two bilinear forms, $(\cdot,\cdot)_k$, and $a_k(\cdot,\cdot)$. While $(\cdot,\cdot)_k$ is just the standard $L^2(\Omega)$-inner product for $k = 0, \ldots, J-1$, at level J, it is a mesh dependent L^2-like inner product defined by

$$(\eta, \mu)_J = \sum_{K \in \mathscr{T}_J} \frac{|K|}{|\partial K|} \int_{\partial K} \eta \, \mu, \tag{9}$$

where $|\cdot|$ denotes the measure. The other bilinear form on $M_k \times M_k$ is defined by

$$a_k(u, v) = \begin{cases} \displaystyle\int_\Omega \nabla u \cdot \nabla v, & k = 0, 1, \cdots, J-1, \\ \displaystyle\int_\Omega \mathcal{Q} u \cdot \mathcal{Q} v, & k = J, \end{cases}$$

where $\mathcal{Q}$ is the previously defined flux-lifting of the hybridized method (on the finest level mesh $\mathscr{T}_J \equiv \mathscr{T}_h$). Hence, defining multilevel operators $A_k : M_k \mapsto M_k$ by

$$(A_k \omega, \varphi)_k = a_k(\omega, \varphi), \qquad \text{for all } \varphi, \ \omega \in M_k,$$

our main goal can be stated as to find an efficient scheme for solving a finest level equation $A_J\lambda_J = b_J$. Note that the forms are non-inherited at the last level. This means in our analysis, we will have to use a multigrid theory general enough to admit non-inherited forms and non-nested spaces.

The main complication in a non-nested setting is the necessity of designing appropriate intergrid transfer operators for moving data back and forth between the multilevel grids. We define the *prolongation operator* $I_k : M_{k-1} \to M_k$, $(k = 2, \ldots, J)$ by

$$I_k v = \begin{cases} v, & \text{for } k < J, \\ v|_{\mathscr{E}_J}, & \text{for } k = J \text{ and } d > 0, \\ \Pi_{M_J}(v|_{\mathscr{E}_J}), & \text{for } k = J \text{ and } d = 0. \end{cases} \tag{10}$$

where $\Pi_{M_J} : L^2(\mathscr{E}_J) \to M_J$ is the $L^2(\mathscr{E}_J)$-orthogonal projection onto M_J. It is important to note that there are many naive choices of intergrid transfer operators that does not work in our application. In Section 5, we shall show numerical experiments with certain "obvious" transfer operators that lead to slow convergence of multigrid. The reverse movement of data, from fine to coarse levels, is achieved through the restriction operator $Q_{k-1} : M_k \to M_{k-1}$, defined by

$$(Q_{k-1}\omega, \varphi)_{k-1} = (\omega, I_k\varphi)_k, \quad \text{for all } \varphi, \omega \in M_{k-1}.$$

The only remaining significant ingredient of the multigrid algorithm is a set of *smoothing operators* $R_k : M_k \mapsto M_k$. The smoother R_k, is chosen to be one of the classical relaxation iterations of Jacobi or Gauss-Seidel. To symmetrize the algorithm, we will also need the the adjoint smoother R_k^t defined by

$$(R_k u, v)_k = (u, R_k^t v)_k \quad \forall u, v \in M_k$$

and the ancillary notation

$$R_k^{(l)} = \begin{cases} R_k & \text{if } l \text{ is odd,} \\ R_k^t & \text{if } l \text{ is even.} \end{cases}$$

The algorithm given below performs m_k pre- and post-smoothings at level k. Our convergence result is under an assumption that the number of smoothings increase in a specific way (detailed in Theorem 3.1) as we proceed to the coarser levels.

Algorithm 3.1 (Variable V-cycle) *Given an initial approximation $u^{(i)} \in M_J$ to the solution of $A_J u = f$, we define the next approximation $u^{(i+1)} \in M_J$ by the iteration*

$$u^{(i+1)} = \mathtt{MG}_J(u^{(i)}, f)$$

where the map $\mathtt{MG}_k(\cdot,\cdot) : M_k \times M_k \mapsto M_k$ is defined recursively as follows.

1. *First, at the coarsest level, set $\mathtt{MG}_1(u, f) = A_1^{-1} f$.*
2. *Next, for $k \geq 2$, define $\mathtt{MG}_k(u^{(i)}, f)$ by the following steps:*

   (a) Set $v^{(0)} = u^{(i)}$.
   (b) (Pre-smoothing) *For $l = 1, 2, \ldots, m_k$,*

$$v^{(l)} = v^{(l-1)} + R_k^{(l+m_k)}(f - A_k v^{(l-1)}).$$

(c) Set residual

$$r_k = f - A_k v^{(m_k)}.$$

(d) (Correction) *Set*

$$q_{k-1} = \mathtt{MG}_{k-1}(0, Q_{k-1} r_k)$$

and

$$w^{(m_k)} = v^{(m_k)} + I_k q_{k-1}$$

(e) (Post-smoothing) *For* $l = m_k + 1, \ldots, 2m_k$,

$$w^{(l)} = w^{(l-1)} + R_k^{(l+m_k)}(f - A_k w^{(l-1)})$$

(f) Finally, define the next iterate by setting

$$\mathtt{MG}_k(u^{(i)}, f) = w^{(2m_k)}.$$

We can use this algorithm as both a linear iteration and as a preconditioner. When using as a linear iteration, we start with an initial guess $u^{(0)}$ and compute successive iterative approximations by $u^{(i+1)} = \mathtt{MG}_J(u^{(i)}, f)$. The iterative error, namely $u - u^{(i)}$, is propagated through an error reducing operator which we denote by $\mathcal{E}_J$, i.e.,

$$u - u^{(i+1)} = \mathcal{E}_J(u - u^{(i)}).$$

It is well known [8] that $\mathcal{E}_J$ is a linear operator admitting a recursive expression. Using the abstract theory of [7, 8], we prove that this iterative error decreases geometrically at a mesh-independent rate, as stated in the next theorem.

Theorem 3.1. *Suppose the number of smoothings, m_k increases as k decreases in such a way that $\beta_0 m_k \le m_{k-1} \le \beta_1 m_k$ for some fixed constants $1 < \beta_0 \le \beta_1$. Assume that Ω is convex. Then there exists a positive $\delta < 1$, independent of the mesh size h_J, such that the error reducing operator of Algorithm 3.1 satisfies*

$$0 \le\ a_J(\mathcal{E}_J u, u)\ \le\ \delta\, a_J(u, u), \qquad \text{for all } u \in M_J.$$

This is the main result of this paper. Its proof is in the next section. The convexity of the domain is assumed so that we can use well known regularity results. Numerical experience indicates that the algorithm converges even when this assumption does not hold. The assumption on the number of smoothings can be easily satisfied, for example, by setting $m_k = 2^{J-k}$, maintaining optimal work count.

We can use Algorithm 3.1 in a preconditioned conjugate gradient iteration. Since the algorithm defines a linear iteration, the operator $B_J : M_J \mapsto M_J$ defined by

$$B_J g = \mathtt{MG}_J(0, g), \qquad \text{for all } g \in M_J$$

is a linear operator. If this is used as a preconditioner for A_J in a conjugate gradient iteration, then the rate of convergence is governed by the condition number $\kappa(B_J A_J)$. Since $\mathcal{E}_J = I - B_J A_J$, Theorem 3.1 shows that $\kappa(B_J A_J)$ is bounded above and below by mesh independent constants. Hence B_J is an optimal preconditioner.

4. Proof of the convergence result

This section is devoted to the proof of Theorem 3.1. We shall use the abstract multigrid theory of [7, 8] which allows the use of non-inherited forms and non-nested spaces. According to this theory, once we verify the following three conditions, the proof of Theorem 3.1 is complete.

Condition 4.1 (Prolongation norm) *For all* $k = 1, \ldots, J$,

$$a_k(I_k v, I_k v) \leq a_{k-1}(v, v), \qquad \forall v \in M_{k-1}.$$

Condition 4.2 (Regularity & Approximation) *There exist* $0 < \alpha \leq 1$ *and* $C > 0$ *such that*

$$a_k((I - I_k P_{k-1})v, v) \leq C \left(\frac{\|A_k v\|_k^2}{\lambda_k} \right)^{\alpha} a_k(v, v)^{1-\alpha}, \quad \forall v \in M_k,\ k = 1, \ldots, J,$$

where $P_{k-1} : M_k \to M_{k-1}$ *is defined by*

$$a_{k-1}(P_{k-1}\omega, \varphi) = a_k(\omega, I_k \varphi), \quad \text{for all } \varphi \in M_{k-1}$$

Condition 4.3 (Smoothing) *There exists* $\omega > 0$ *such that*

$$\omega \frac{\|v\|_k}{\lambda_k} \leq (\tilde{R}_k v, v), \qquad \forall v \in M_k,\ k = 1, \ldots, J,$$

where $\tilde{R}_k = R_k + R_k^t - R_k A_k R_k^t$ *and* λ_k *is the eigenvalue of* A_k *with maximal norm.*

The remainder of this section is divided into three subsections, each devoted to the verification of one of the above conditions.

4.1. Verification of Condition 4.1

This condition limits growth in prolongation norms. In our application, for $k = 1, \ldots, J-1$, the prolongation I_k is the identity and $a_{k-1}(\cdot, \cdot) = a_k(\cdot, \cdot)$, so the condition obviously holds. Hence, it only remains to consider the case $k = J$. This follows from the next lemma.

Lemma 4.1. *For any* $v_{J-1} \in M_{J-1}$,

$$\mathcal{Q}(I_J v_{J-1}) = -\nabla v_{J-1}. \tag{11}$$

Proof. Let the divergence free subspace of the Raviart-Thomas space be denoted by

$$R_d^0(K) = \{\boldsymbol{q} \in R_d(K) : \quad \operatorname{div}(\boldsymbol{q}|_K) = 0\} \tag{12}$$

for any $K \in \mathscr{T}_J$. Then, by the definition of the flux-lifting $\mathcal{Q}(\cdot)$,

$$\int_K \mathcal{Q}(I_J v_{J-1}) \cdot \boldsymbol{r} = -\int_{\partial K} (I_J v_{J-1}) \boldsymbol{r} \cdot \boldsymbol{n}, \qquad \forall \boldsymbol{r} \in R_d^0(K).$$

First consider the case $d = 0$. Then, by (10), the right hand side above can be rewritten as

$$\begin{aligned} -\int_{\partial K} (I_J v_{J-1}) \boldsymbol{r} \cdot \boldsymbol{n} &= -\int_{\partial K} (\Pi_{M_J} v_{J-1}) \boldsymbol{r} \cdot \boldsymbol{n} \\ &= -\int_{\partial K} v_{J-1} \boldsymbol{r} \cdot \boldsymbol{n} = -\int_K (\nabla v_{J-1}) \cdot \boldsymbol{r}, \end{aligned}$$

where the last two equalities follow because $\boldsymbol{r}\cdot\boldsymbol{n}$ is piecewise constant when $d=0$, and by integration by parts, respectively. Since both $\mathcal{Q}(\lambda_J)|_K$ and $\nabla v_{J-1}|_K$ are in $R_d^0(K)$, the above proves the lemma in the $d=0$ case.

When $d>0$, selecting the appropriate case in the definition of I_k in (10), we have

$$-\int_{\partial K}(I_J v_{J-1})\boldsymbol{r}\cdot\boldsymbol{n} = -\int_{\partial K} v_{J-1}\boldsymbol{r}\cdot\boldsymbol{n} \;= -\int_K (\nabla v_{J-1})\cdot\boldsymbol{r},$$

so the proof can be completed as before. □

The verification of Condition 4.1 is now completed by observing that because of Lemma 4.1,

$$a_J(I_J v, I_J v) = (\mathcal{Q}(I_J v), \mathcal{Q}(I_J v)) = (\nabla v_{J-1}, \nabla v_{J-1}) = a_{J-1}(v_{J-1}, v_{J-1}).$$

Here $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ inner product and is not to be confused with $(\cdot,\cdot)_J$ defined in (9).

4.2. Verification of Condition 4.2

Inequalities like that of Condition 4.2 typically follow as a consequence of some regularity results for the underlying boundary value problem, combined with the approximation properties of the finite element spaces. It is well-known that Condition 4.2 holds for $k = 0, 1\ldots, J-1$ [4]. (This is part of the standard full regularity based proofs of multigrid convergence for the continuous Galerkin method [2, 3, 8].) So we only need to verify Condition 4.2 with $k=J$.

For this, we need a number of intermediate lemmas that establish properties of various local operators. Let us begin with the local lifting operator $\mathcal{U}(\cdot)$ defined earlier.

Lemma 4.2. *For all $w\in P_1(K)$, we have*

$$\mathcal{U}(I_J w) = \Pi_{T_h} w$$

where Π_{T_h} is the $L^2(\Omega)$-orthogonal projection onto T_h.

Proof. Given $w\in P_1(K)$, by the definition of the lifting operators in (5), we have

$$\int_K \mathcal{Q}(I_J w)\cdot\boldsymbol{r} - \int_K \mathcal{U}(I_J w)\,\mathrm{div}\boldsymbol{r} = -\int_{\partial K} I_J w\;(\boldsymbol{r}\cdot\boldsymbol{n})$$

for all $\boldsymbol{r}\in R_d(K)$. On the right hand side above, we can replace $I_J w$ by w if $d>0$. We can also do this if $d=0$, because in this case $\boldsymbol{r}\cdot\boldsymbol{n}$ takes a constant value on each edge. Therefore, by an integration by parts formula, we obtain

$$\int_K \mathcal{Q}(I_J w)\cdot\boldsymbol{r} - \int_K \mathcal{U}(I_J w)\mathrm{div}\boldsymbol{r} = -\int_K \nabla w\cdot\boldsymbol{r} - \int_K w\,\mathrm{div}\boldsymbol{r}$$

Since $\mathcal{Q}(I_J w) = -\nabla w$ by Lemma 4.1, this implies

$$\int_K (\mathcal{U}(I_J w) - w)\,\mathrm{div}\boldsymbol{r} = 0 \qquad \forall \boldsymbol{r}\in R_d(K).$$

The lemma now follows, since $\mathrm{div} : R_d(K)\to P_d(K)$ is a surjection. □

Next, we need to define a new local operator that maps a pair of interior and boundary functions into one function. Let $L_d(K) = \{p \in P_{d+3}(K) : p|_e \in P_{d_+}(e), \forall \text{ edge } e \text{ of K}\}$, where

$$d_+ = \begin{cases} d+1, & \text{if } d \text{ is even,} \\ d+2, & \text{if } d \text{ is odd.} \end{cases}$$

Suppose we are given $p \in L^2(K)$ and $\lambda \in L^2(\partial K)$. Consider a function $\psi(p,\lambda) \in L_d(K)$ that satisfies

$$\begin{aligned} \int_K \psi(p,\lambda)\, s &= \int_K p\, s, \quad \forall s \in P_d(K), \text{ and} \\ \int_e \psi(p,\lambda)\, \mu &= \int_e \lambda\mu, \quad \forall \mu \in P_{d_+-1}(e), \end{aligned} \tag{13}$$

for all the three edges e of K. That such a ψ is unique is proved next. As usual, when performing standard scaling arguments, we obtain constants that depend on the shape regularity of the mesh, namely on a fixed constant Υ which is the maximum of $\operatorname{diam}(K)/\rho_K$ over all elements K, where ρ_K denotes the diameter of the largest ball inscribed in K.

Lemma 4.3. *There is a unique $\psi(p,\lambda)$ in $L_d(K)$ satisfying (13). Furthermore, there are constants C_1 and C_2 depending only on the shape regularity constant Υ such that*

$$C_1\|\psi(p,\lambda)\|_{L^2(K)} \le \|p\|_{L^2(K)} + |\partial K|^{1/2}\|\lambda\|_{L^2(\partial K)} \le C_2\|\psi(p,\lambda)\|_{L^2(K)}$$

for all p in $P_d(K)$ and all λ such that $\lambda|_e$ is in $P_d(e)$ for all three edges e of K.

Proof. First, we check if (13) forms a square system for $\psi(p,\lambda)$. Indeed, the number of equations in the system (13) equals

$$\dim(P_d(K)) + 3\dim(P_{d_+-1}(e)) = \frac{1}{2}(d+1)(d+2) + 3d_+. \tag{14}$$

On the other hand, the number of degrees of freedom of $L_d(K)$ can be counted by adding together the dimension of $P_1(K)$ (equaling 3), the dimension of the space of all edge bubbles of $L_d(K)$ (equaling $3(d_+ - 1)$), and the dimension of interior bubbles of $L_d(K)$ (equaling $(d+1)(d+2)/2$). Thus,

$$\dim(L_d(K)) = 3 + 3(d_+ - 1) + \frac{1}{2}(d+1)(d+2),$$

which simplifies to the same number as in (14). Thus (13) is a square system.

To prove that there is a unique $\psi(p,\lambda)$ satisfying (13), it now suffices to show that if p and λ vanish, the only solution of (13) is trivial. To this end, consider a ψ in $L_d(K)$ satisfying

$$\int_K \psi s = 0, \quad \forall s \in P_d(K) \tag{15}$$

$$\int_e \psi\mu = 0, \quad \forall e, \forall \mu \in P_{d_+-1}(e). \tag{16}$$

The last equation (16) implies that on each edge e, $\psi|_e$ is a polynomial on $P_{d_+}(e)$ that is orthogonal to all $P_{d_+-1}(e)$. Hence $\psi|_e$ must be the Legendre polynomial of degree d_+. No matter what d is, d_+ is always odd, hence $\psi|_e$ is an odd function on the edge e. Since this

holds for all three edges, and since ψ must be continuous on ∂K, we conclude that ψ vanishes on ∂K.

Since $\psi \in L_d$ vanishes on all the three edges, it must have the form

$$\psi = \lambda_1\lambda_2\lambda_3 p_d, \qquad \text{for some } p_d \in P_d(K)$$

where λ_i are the barycentric coordinates of K. Hence (15) implies

$$\int_K (\lambda_1\lambda_2\lambda_3) p_d \, s = 0, \qquad \forall s \in P_d(K).$$

Therefore, $p_d \equiv 0$, and consequently, $\psi \equiv 0$. This proves the unique solvability of (13).

The norm estimate of the lemma follows because if p and λ are as in the statement of the lemma, then $\psi(p,\lambda) = 0$ if and only if $p = 0$ and $\lambda = 0$. Thus, on a fixed reference element $\hat{K}$, the norms $\left(\|p\|^2_{L^2(\hat{K})} + \|\lambda\|^2_{L^2(\partial\hat{K})}\right)^{1/2}$ and $\|\psi(p,\lambda)\|_{L^2(\hat{K})}$ are equivalent. The stated norm estimate then follows by a scaling argument mapping $\hat{K}$ to K. □

Using the above defined element space $L_d(K)$ on each mesh element, we can define a new lifting of λ from the element boundaries into the element interiors by

$$S\lambda = \psi(\mathcal{U}\lambda, \lambda). \tag{17}$$

The next lemma establishes a few properties of S that we need. In its statement, and in the remainder, $\|\lambda\|_a$ denotes the "energy"-like norm on the finest level, i.e.,

$$\|\lambda\|_a^2 = a_J(\lambda,\lambda) \qquad \text{for all } \lambda \in M_J,$$

and $\|\lambda\|_J$ is the norm defined in (9).

Lemma 4.4. *For any λ in M_J, the following statements hold:*

$$C_1\|\lambda\|_a^2 \geq \sum_{K\in\mathscr{T}_J} \left|S\lambda\right|^2_{H^1(K)} \tag{18}$$

$$C_2\|\lambda\|_J \leq \|S\lambda\|_{L^2(\Omega)} \leq C_3\|\lambda\|_J \tag{19}$$

$$\mathcal{U}\lambda = \Pi_{T_h}(S\lambda). \tag{20}$$

Here C_i's are mesh independent constants.

Proof. To prove (18), first observe that if λ takes a constant value κ on the boundary of some mesh element ∂K, then $S\lambda$ takes the same constant value κ on K. (This follows from Lemma 4.2.) Hence, for any λ, we have

$$\nabla(S\bar{\lambda}\big|_K) = 0, \qquad \text{where} \quad \bar{\lambda} = \frac{1}{|\partial K|}\left(\int_{\partial K}\lambda\right).$$

Therefore,

$$\begin{aligned}
\|\nabla(S\lambda)\|_{L^2(K)} &= \|\nabla S\lambda - \nabla S\bar{\lambda}\|_{L^2(K)}, \\
&\leq Ch^{-1}\|S(\lambda-\bar{\lambda})\|_{L^2(K)} && \text{by inverse inequality} \\
&\leq Ch^{-1}(\|\mathcal{U}(\lambda-\bar{\lambda})\|_{L^2(K)} + |\partial K|^{1/2}\|\lambda-\bar{\lambda}\|_{L^2(\partial K)}) && \text{by Lemma 4.3} \\
&\leq Ch^{-1}(|\partial K|^{1/2}\|\lambda-\bar{\lambda}\|_{L^2(\partial K)}) && \text{by [15, Lemma 3.3]} \\
&\leq C|\partial K|^{-1/2}\|\lambda-\bar{\lambda}\|_{L^2(\partial K)}.
\end{aligned}$$

Summing over all elements and using a norm equivalence proved in [19, Theorem 2.2], we get

$$\sum_{K\in\mathscr{T}_J} \|\nabla(S\lambda)\|^2_{L^2(K)} \le C \sum_{K\in\mathscr{T}_J} |\partial K|^{-1}\|\lambda-\bar{\lambda}\|^2_{L^2(\partial K)} \le Ca_h(\lambda,\lambda)$$

which proves (18).

The proof of (19) is a straightforward consequence of Lemma 4.3.

The identity (20) is obvious from (13) and (17). □

We need one more intermediate map before we can give our proof of Condition 4.2. To describe this map, first we define a ϕ_λ in M_J for every λ in M_J by

$$(S\phi_\lambda, S\mu) = (\mathcal{Q}\lambda, \mathcal{Q}\mu), \quad \forall \mu \in M_J. \tag{21}$$

This equation is uniquely solvable for ϕ_λ in M_J, because if the right-hand side is zero, then $S\phi_\lambda = 0$, so $\phi_\lambda = 0$ by the estimate (19) of Lemma 4.4. Next, let $f_\lambda = \mathcal{U}\phi_\lambda$. The map we use in the later proof is a map from M_J into M_J, which for notational simplicity, we denote by

$$\lambda \longmapsto \tilde{\lambda}$$

where $\tilde{\lambda} \in M_J$ is the unique solution of the equation

$$a(\tilde{\lambda}, \mu) = (f_\lambda, \mathcal{U}\mu), \quad \forall \mu \in M_J. \tag{22}$$

The following lemma reveals the relationship between $\lambda, \tilde{\lambda}$ and ϕ_λ.

Lemma 4.5. *Let $\lambda, \tilde{\lambda}$ and ϕ_λ be defined as above. Then*

$$\|S\phi_\lambda\|_{L^2(\Omega)} \le C\|A_J\lambda\|_J \tag{23}$$

$$\|\lambda - \tilde{\lambda}\|_a \le Ch_J\|A_J\lambda\|_J. \tag{24}$$

Proof. To prove (23),

$$\begin{aligned}
\|S\phi_\lambda\|_{L^2(\Omega)} &= \sup_{\mu\in M_J} \frac{(S\phi_\lambda, S\mu)}{\|S\mu\|_{L^2(\Omega)}} = \sup_{\mu\in M_J} \frac{a_J(\lambda,\mu)}{\|S\mu\|_{L^2(\Omega)}} && \text{by (21)} \\
&= \sup_{\mu\in M_J} \frac{(A_J\lambda,\mu)_J}{\|S\mu\|_{L^2(\Omega)}} \le C \sup_{\mu\in M_J} \frac{(A_J\lambda,\mu)_J}{\|\mu\|_J} && \text{by Lemma 4.4, (19)} \\
&\le C\|A_J\lambda\|_J.
\end{aligned}$$

To prove (24), note that

$$\begin{aligned}
a(\lambda,\mu) &= (S\phi_\lambda, S\mu) \\
a(\tilde{\lambda},\mu) &= (\Pi_{T_h} S\phi_\lambda, S\mu),
\end{aligned}$$

where the last identity follows because in (22), the right hand side function is $f_\lambda = \mathcal{U}\phi_\lambda = \Pi_{T_h}(S\phi_\lambda)$ by Lemma 4.4, (20). Subtracting, and setting $\mu = \lambda - \tilde{\lambda}$, we get

$$\|\lambda-\tilde{\lambda}\|^2_a = ((I-\Pi_{T_h})S\phi_\lambda, S(\lambda-\tilde{\lambda})) = ((I-\Pi_{T_h})S\phi_\lambda, (I-\Pi_{T_h})S(\lambda-\tilde{\lambda})).$$

Using the Friedrichs estimate $\|u - \Pi_{T_h} u\|_{L^2(K)} \le C h_J |u|_{H^1(K)}$, we get

$$\begin{aligned}
\|\lambda - \tilde{\lambda}\|_a^2 &\le \Big(\sum_{K \in \mathscr{T}_J} C h_J^2 |S\phi_\lambda|^2_{H^1(K)} \Big)^{1/2} \Big(\sum_{K \in \mathscr{T}_J} C h_J^2 |S(\lambda - \tilde{\lambda})|^2_{H^1(K)} \Big)^{1/2} \\
&\le \Big(\sum_{K \in \mathscr{T}_J} C \|S\phi_\lambda\|^2_{L^2(K)} \Big)^{1/2} \Big(\sum_{K \in \mathscr{T}_J} C h_J^2 |S(\lambda - \tilde{\lambda})|^2_{H^1(K)} \Big)^{1/2} \\
&\le C \|A_J \lambda\|_J \Big(\sum_{K \in \mathscr{T}_J} C h_J^2 |S(\lambda - \tilde{\lambda})|^2_{H^1(K)} \Big)^{1/2}
\end{aligned}$$

by an inverse inequality and (23). Finally, applying the estimate (18) of Lemma 4.4 on the last term above, and canceling the common factor, we obtain (24). □

With these preparations, we can now finish the verification of Condition 4.2.

Proof of Condition 4.2 with $\alpha = 1$. Let λ be in M_J. We need to estimate the quantity

$$a_J((I - I_J P_{J-1})\lambda, \lambda) = \int_\Omega \mathcal{Q}(\lambda - I_J P_{J-1}\lambda) \cdot \mathcal{Q}\lambda$$

To estimate this, it will be useful to make the following preliminary observations: First, by Lemma 4.1,

$$\mathcal{Q}(\lambda - I_J P_{J-1}\lambda) = \mathcal{Q}(\lambda) + \nabla(P_{J-1}\lambda). \tag{25}$$

Second, the expression on the right hand side above satisfies

$$(\mathcal{Q}\lambda + \nabla(P_{J-1}\lambda), \nabla v_{J-1})_\Omega = 0, \qquad \forall v_{J-1} \in M_{J-1}, \tag{26}$$

because

$$\begin{aligned}
(\nabla(P_{J-1}\lambda), \nabla v_{J-1}) &= a_{J-1}(P_{J-1}\lambda, v_{J-1}) \\
&= a_J(\lambda, I_J v_{J-1}) \\
&= (\mathcal{Q}\lambda, \mathcal{Q}(I_J v_{J-1})) \\
&= -(\mathcal{Q}\lambda, \nabla v_{J-1}).
\end{aligned}$$

Thus,

$$a_J((I - I_J P_{J-1})\lambda, \lambda) = \|\mathcal{Q}\lambda + \nabla(P_{J-1}\lambda)\|^2, \tag{27}$$

and it suffices to estimate $\mathcal{Q}\lambda + \nabla(P_{J-1}\lambda)$. Here, and in the remainder, we use $\|\cdot\|$ (without any subscripts) as well as $\|\cdot\|_{L^2(\Omega)}$ to denote the $L^2(\Omega)$-norm.

To begin the estimation, we split $\mathcal{Q}\lambda + \nabla(P_{J-1}\lambda)$ into many terms, labeling each term as follows:

$$\begin{aligned}
\mathcal{Q}\lambda + \nabla(P_{J-1}\lambda) = {} & \mathcal{Q}(\lambda - \tilde{\lambda}) && \cdots\cdots\cdots \text{(term A)} \\
& + \mathcal{Q}\tilde{\lambda} - \mathcal{Q}(\Pi_{M_J}\tilde{u}) && \cdots\cdots\cdots \text{(term B)} \\
& + \mathcal{Q}(\Pi_{M_J}\tilde{u}) - (-\nabla\tilde{u}) && \cdots\cdots\cdots \text{(term C)} \\
& + \nabla P_{J-1}\tilde{\lambda} - \nabla\tilde{u} && \cdots\cdots\cdots \text{(term D)} \\
& + \nabla P_{J-1}(\lambda - \tilde{\lambda}), && \cdots\cdots\cdots \text{(term E)}
\end{aligned} \tag{28}$$

where $\tilde{\lambda}$ is as defined in (22) and $\tilde{u}$ is the unique function in $H^1_0(\Omega)$ that solves

$$(\nabla \tilde{u}, \nabla v) = (f_\lambda, v), \quad \forall v \in H^1_0(\Omega).$$

Note that since we have assumed a convex domain,

$$\|\tilde{u}\|_{H^2(\Omega)} \leq C \|f_\lambda\|_{L^2(\Omega)}. \tag{29}$$

by a well known regularity theorem [20].

The first term can be estimated by

$$\|(\text{term A})\| = \|\mathcal{Q}(\lambda - \tilde{\lambda})\| = \|\lambda - \tilde{\lambda}\|_a \leq C h_J \|A_J \lambda\|_J$$

by the inequality (24) of Lemma 4.5. For the next term, first observe that due to the characterization of Lagrange multipliers given by Theorem 2.1, $\tilde{\lambda}$ is the hybridized mixed method approximation to $\tilde{u}$. Hence by a previously established Lagrange multiplier error estimate [15, Theorem 3.1],

$$\|\mathcal{Q}\tilde{\lambda} - \mathcal{Q}(\Pi_{M_J} \tilde{u})\| \leq \|\boldsymbol{q} - \Pi_R \boldsymbol{q}\| \tag{30}$$

where $\boldsymbol{q} = -\nabla \tilde{u}$ and $\Pi_R \boldsymbol{q}$ is the Raviart-Thomas interpolant of $\boldsymbol{q}$. By the standard error estimates for this interpolant, we immediately find that

$$\begin{aligned}
\|(\text{term B})\| = \|\mathcal{Q}\tilde{\lambda} - \mathcal{Q}(\Pi_{M_J} \tilde{u})\| &\leq C h_J |\tilde{u}|_{H^2(\Omega)} \\
&\leq C h_J \|f_\lambda\|_{L^2(\Omega)} \qquad \text{by (29)} \\
&= C h_J \|\Pi_{T_h}(S\phi_\lambda)\|_{L^2(\Omega)} \leq C h_J \|S\phi_\lambda\|_{L^2(\Omega)} \\
&\leq C h_J \|A_J \lambda\|_J,
\end{aligned}$$

using the estimate (23) of Lemma 4.5.

We proceed to analyze the next term. For this, recall the divergence free subspace $R^0_d(K)$ defined in (12). By the definition of $\mathcal{Q}(\cdot)$,

$$\begin{aligned}
\int_K \mathcal{Q}(\Pi_{M_J} \tilde{u}) \cdot \boldsymbol{r} &= -\int_{\partial K} (\Pi_{M_J} \tilde{u}) \boldsymbol{r} \cdot \boldsymbol{n} = -\int_{\partial K} \tilde{u}\, \boldsymbol{r} \cdot \boldsymbol{n} \\
&= -\int_K \nabla \tilde{u} \cdot \boldsymbol{r}
\end{aligned} \tag{31}$$

Now, suppose $\tilde{u}_{J-1} \in M_{J-1}$ is the exact solution of

$$(\nabla \tilde{u}_{J-1}, \nabla v) = (f_\lambda, v),\ \forall v \in M_{J-1}.$$

Then, on any mesh element K, its flux $-\nabla \tilde{u}_{J-1}|_K$ is constant, and therefore in $R^0_d(K)$. Putting $\boldsymbol{r} = \mathcal{Q}(\Pi_{M_J} \tilde{u}) + \nabla \tilde{u}_{J-1}|_K$ in (31), we have

$$\big(\mathcal{Q}(\Pi_{M_J} \tilde{u}) + \nabla \tilde{u},\ \mathcal{Q}(\Pi_{M_J} \tilde{u}) + \nabla u_{J-1}\big)_K = 0,$$

or in other words,

$$\|\mathcal{Q}(\Pi_{M_J} \tilde{u}) + \nabla \tilde{u}\|^2_{L^2(K)} = \big(\mathcal{Q}(\Pi_{M_J} \tilde{u}) + \nabla \tilde{u}, \nabla \tilde{u} - \nabla \tilde{u}_{J-1}\big)_K.$$

By Cauchy-Schwarz inequality,

$$\begin{aligned}
\|(\text{term C})\| &= \|\mathcal{Q}(\Pi_{M_J} \tilde{u}) + \nabla \tilde{u}\|_{L^2(\Omega)} \\
&\leq \|\nabla(\tilde{u} - \tilde{u}_{J-1})\|_{L^2(\Omega)} \leq C h_J |\tilde{u}|_{H^2(\Omega)} \leq C h_J \|f_\lambda\|_{L^2(\Omega)} \\
&\leq C h_J \|A_J \lambda\|_J,
\end{aligned} \tag{32}$$

where we have used the standard error estimate for conforming linear finite elements (since $\tilde{u}_{J-1}$ is the conforming linear finite element approximation of $\tilde{u}$), the regularity estimate (29), and (23) of Lemma 4.5.

For (term D), we will first show that $P_{J-1}\tilde{\lambda}$ coincides with the $\tilde{u}_{J-1}$ defined above. Indeed, for all $w_{J-1} \in M_{J-1}$, we have

$$\begin{aligned}
(\nabla P_{J-1}\tilde{\lambda}, \nabla w_{J-1}) &= -(\mathcal{Q}\tilde{\lambda}, \nabla w_{J-1}) && \text{by (26)} \\
&= (\mathcal{Q}\tilde{\lambda}, \mathcal{Q}(I_J w_{J-1})) && \text{by Lemma 4.1} \\
&= (f_\lambda, \mathcal{U}(I_J w_{J-1})) && \text{by (22) and Theorem 2.1} \\
&= (f_\lambda, w_{J-1}) && \text{by Lemma 4.2.}
\end{aligned}$$

Thus $P_{J-1}\tilde{\lambda}$ and $\tilde{u}_{J-1}$ satisfy the same equations in M_{J-1} and must coincide. Therefore

$$\|(\text{term D})\| = \|\nabla(P_{J-1}\widetilde{\lambda} - \tilde{u})\| = \|\nabla(\tilde{u} - \tilde{u}_{J-1})\| \leq Ch_J \|A_J\lambda\|_J,$$

by the same arguments as in (32).

For the final term, we first note that if we choose $\lambda = \mu$ and $v_{J-1} = P_{J-1}\mu$ in (26), then we have

$$\|\nabla P_{J-1}\mu\|^2 = -(\mathcal{Q}\mu, \nabla P_{J-1}\mu),$$

and hence, by Cauchy-Schwarz inequality,

$$\|\nabla P_{J-1}\mu\| \leq \|\mathcal{Q}\mu\|, \quad \forall \mu \in M_{J-1}$$

Therefore

$$\|(\text{term E})\| = \|\nabla P_{J-1}(\lambda - \tilde{\lambda})\| \leq \|\mathcal{Q}(\lambda - \tilde{\lambda})\| \leq Ch_J \|A_J\lambda\|_J$$

by the estimate (24) of Lemma 4.5.

Returning to (27) and combining the estimates for each of the terms above, we obtain

$$\begin{aligned}
a_J((I - I_J P_{J-1})\lambda, \lambda) &= \|\mathcal{Q}((I - I_J P_{J-1})\lambda)\|^2 \\
&\leq Ch_J^2 \, \|A_J\lambda\|_J^2
\end{aligned}$$

By [19, Theorem 2.3], we know that $\lambda_J \leq Ch_J^{-2}$. Hence the above inequality proves Condition 4.2 with $\alpha = 1$. □

4.3. Verification of Condition 4.3

We need to verify this smoothing condition for the Jacobi and the Gauss-Seidel smoothers. Again, for all the levels $k = 1, \ldots, J-1$, the result is standard [8]. For the highest level $k = J$, the arguments are also fairly standard. Nonetheless, we will sketch the proof for this case now.

Both the Jacobi and Gauss-Seidel iterations, which are well known classical iterations, can be rewritten using the modern "subspace decomposition" framework. To display this decomposition for the finest level space M_J, let ϕ_J^i, $i = 1, 2, \ldots, N_J$ denote a local basis for M_J with the property that each ϕ_J^i is supported only on one mesh edge. Further, let $M_{J,i} = \text{span}\{\phi_J^i\}$. Then the subspace decomposition of M_J is

$$M_J = \sum_{i=1}^{N_J} M_{J,i}$$

The Jacobi and Gauss-Seidel operators can be written in terms of local operators on these subspaces. Define $A_{J,i} : M_{J,i} \to M_{J,i}$ by

$$(A_{J,i}u, v)_J = a_J(u, v) \quad \text{for all } u, v \in M_{J,i}.$$

Let $Q_{J,i} : M_J \to M_{J,i}$ and $P_{J,i} : M_J \to M_{J,i}$ be defined by

$$\begin{aligned} (Q_{J,i}u, v)_J &= (u, v)_J && \text{for all } u \in M_J,\ v \in M_{J,i}, \\ (P_{J,i}u, v)_J &= a_J(u, v) && \text{for all } u \in M_J,\ v \in M_{J,i}. \end{aligned}$$

Then the operator R_J defining the Jacobi iteration on each M_J is defined as $R_J = \gamma \mathcal{J}_J$, where γ is a scaling parameter and

$$\mathcal{J}_J = \sum_{i=1}^{N_J} A_{J,i}^{-1} Q_{J,i}. \tag{33}$$

The Gauss-Seidel operator on each M_J is defined as

$$\mathcal{G}_J = \big(I - (I - P_{J,N_J})(I - P_{J,N_{J-1}}) \cdots (I - P_{J,1})\big) A_J^{-1}. \tag{34}$$

On the remaining levels, smoothers $\gamma \mathcal{J}_k$ and $\mathcal{G}_k$ can be written out using the standard subspace decompositions of the conforming finite element spaces.

We begin with a simple lemma on the stability of the decomposition in the mesh dependent L^2-like norms on $\mathscr{E}_J$.

Lemma 4.6. *For any set of scalar values c_i,*

$$\sum_{i=1}^{N_J} c_i^2 \|\phi_J^i\|_J^2 \le C \left\| \sum_{i=1}^{N_J} c_i \phi_J^i \right\|_J^2.$$

PROOF. If $v = \sum_{i=1}^{N_J} c_i \phi_J^i \in M_J$, the quantity $\|v\|_J^2$, appearing on the right hand side of the estimate of the lemma, can be evaluated by summing over contributions from each mesh edge. There are $d+1$ basis functions supported on an edge e, which we denote by $\phi_J^{i_1}, \phi_J^{i_2}, \ldots \phi_J^{i_{d+1}}$. By a scaling argument, it is clear that there is a constant κ_d depending only on d, but not on the edge length, such that

$$\sum_{\ell=1}^{d+1} c_{i_\ell}^2 \|\phi_J^{i_\ell}\|_{L^2(e)}^2 \le \kappa_d \left\| \sum_{\ell=1}^{d+1} c_{i_\ell} \phi_J^{i_\ell} \right\|_{L^2(e)}^2.$$

Summing over all edges, we get the result. □

Lemma 4.7. *For all v in M_J,*

$$(\mathcal{J}_J^{-1} v, v) \le C \lambda_J \|v\|_J^2.$$

PROOF. Since $\mathcal{J}_k$ is an additive operator of the form in (33), by a well known lemma on additive

operators [8], splitting $v = \sum_{i=1}^{N_k} c_i \phi_J^i$,

$$\begin{aligned}(\mathcal{J}_J^{-1} v, v) &= \sum_{i=1}^{N_J} a_J(c_i\phi_J^i, c_i\phi_J^i) = \sum_{i=1}^{N_J} c_i^2 (A_J\phi_J^i, \phi_J^i)_J \\ &\leq \sum_{i=1}^{N_J} c_i^2 \lambda_J (\phi_J^i, \phi_J^i)_J = \lambda_J \left(\sum_{i=1}^{N_J} c_i^2 \|\phi_J^i\|_J^2 \right) \\ &\leq C\lambda_J \| \sum_{i=1}^{N_J} c_i \phi_J^i \|_J^2 = C\lambda_J \|v\|_J^2,\end{aligned}$$

where we have used Lemma 4.6. □

From the above lemma, the smoothing conditions can be verified by standard arguments. Indeed, the only other ingredient needed is an inequality of the form

$$\sum_{j=1}^{N_J} \sum_{l=1}^{N_J} |a_J(v_j, w_l)| \leq \beta \left(\sum_{j=1}^{N_J} a_J(v_j, v_j) \right)^{1/2} \left(\sum_{l=1}^{N_J} a_J(w_l, w_l) \right)^{1/2}$$

for all $v_j \in M_{J,j}, w_l \in M_{J,l}$, with some mesh independent constant β. This is often known as a consequence of "limited interaction" of basis functions and is easily verified in our application. Using this result, standard arguments prove [5, 6, 8] the following lemma:

Lemma 4.8. *Choose the scaling parameter such that* $0 < \gamma < \frac{2}{\beta}$. *Then*

1. *if* $R_J = \gamma \mathcal{J}_J$, *then*
$$(\tilde{R}_J^{-1} v, v)_J \leq \gamma^{-1} (\mathcal{J}_J^{-1} v, v)_J, \quad \forall v \in M_J.$$
2. *if* $R_J = \mathcal{G}_J$, *then*
$$(\tilde{R}_J^{-1} v, v) \leq \beta (\mathcal{J}_J^{-1} v, v), \quad \forall v \in M_J.$$

To complete the verification of Condition 4.3 note that

$$\omega \frac{\|v\|_k}{\lambda_k} \leq (\tilde{R}_k v, v) \qquad \forall v \in M_k,$$

holds if and only if

$$(\tilde{R}_k^{-1} v, v) \leq \frac{\lambda_k}{\omega} \|v\|_k^2, \qquad \forall v \in M_k.$$

For $k = J$, the latter inequality follows by combining the estimates of Lemma 4.8 with Lemma 4.7. For the remaining k, the estimate is standard for point Jacobi and Gauss-Seidel operators. Thus Condition 4.3 is verified for all k.

5. Numerical experiments

In this section we report our numerical experiments. We show numerical examples to illustrate the efficacy of our multigrid algorithm. We also show numerical experiments showing the failure of certain naive intergrid transfer operators.

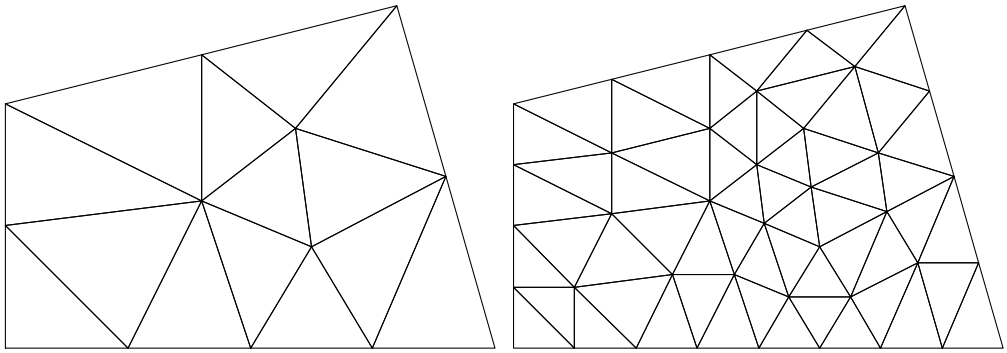

Figure 2. Refining mesh: initial mesh on the left, and refined mesh on the right. Corner coordinators in the initial mesh are $(0,0), (1,0), (0.8,0.7)$ and $(0,0.5)$.

For the first experiment that we shall now describe, we started with a coarse mesh $\mathscr{T}_1$ generated by the public domain meshing software TRIANGLE [24], and then produced a sequence of refinements $\mathscr{T}_2, \mathscr{T}_3, \ldots, \mathscr{T}_J$ by connecting the midpoints of edges, as explained before. The domain and the first two meshes are shown in Figure 2. Suppose we need to solve the Dirichlet problem (1) on the finest mesh level $\mathscr{T}_J$ for various choices of J. The exact solution is $u(x,y) = \sin(x)\, e^{y/2}$. This problem requires a nonzero Dirichlet boundary condition $u = g$ on $\partial\Omega$, which entails the addition of the term $-\int_{\partial\Omega} g\, \boldsymbol{v}\cdot\boldsymbol{n}$ on the right hand side of (4a). But the multigrid algorithm is unaffected.

We use two different algorithms to solve the system $Ax = b$, namely the Conjugate Gradient (CG) iteration, and the variable V-cycle (MG) of Algorithm 3.1 with $m_k = 2^{J-k}$. All experiments are done with the lowest order method, i.e., $d = 0$. We start with the zero function as the initial iterate and stop the iterations when the initial error is reduced by a factor of 10^{-8}. We list the results in Table I. All experiments are run on Intel Core Duo processor (CPU @1.73GHz, 512 Mb RAM). After solving the system for λ, we recover both u and $\boldsymbol{q}$ as described in Section 2. Three different kinds of discretization errors are reported in Table II. They show the convergence of finite element error in accordance with the known theoretical results [1, 15].

As can be seen from the last column of the table, when the size of the matrix increases by a factor of about 4 (which happens when h is halved), the number of multigrid iterations as well as the CPU time in seconds also increases by a factor of 4. This indicates that our multigrid algorithm indeed gives an iterative process with the asymptotically optimal $O(N)$ cost, where N is the number of unknowns. At the same time, the cost increases by a factor of around 8 for CG , each time the mesh size is halved. This clearly demonstrates the benefits of the multigrid algorithm. Also notice that the number of MG iterations seems bounded even as the matrix size gets very large. This is in accordance with the conclusion of Theorem 3.1. In other words, the error reduction factor seems to be independent of mesh size, which is in accordance with the conclusion of Theorem 3.1.

Next, we present an example designed to check if the sufficient condition that Ω is convex (in Theorem 3.1) is necessary. We repeated the experiments with the domain as shown in Figure 3. Table III gives the experimental data. The numbers of multigrid iteration still seems to remain bounded. We conclude that our multigrid algorithm can be effective even when Ω

Size	CG		MG	
	Iterations	CPU secs	Iterations	CPU secs
74	48	0.00	20	0.01
316	101	0.02	26	0.02
1304	207	0.19	31	0.10
5296	418	1.64	33	0.39
21344	833	17.66	34	2.18
85696	1658	148.03	34	9.6
343424	3283	1209.65	34	40.68
1374976	6503	10190.00	34	163.08
5502464	*	*	34	668.63
22014976	*	*	34	2607.65

Table I. Performance comparison between unpreconditioned conjugate gradient method and the multigrid method. (Entries marked * indicates unavailable data due to excessive computational time.)

h	$\|Pu - \lambda_h\|_A$	$\|u - u_h\|_{L^2}$	$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2}$
1	0.04194242	0.05766834	0.06060012
1/2	0.02699368	0.02876214	0.03081513
1/4	0.01099286	0.01437199	0.01552051
1/8	0.00552480	0.00718485	0.00778076
1/16	0.00276706	0.00359228	0.00389374
1/32	0.00138425	0.00179612	0.00194739
1/64	0.00069224	0.00089806	0.00097377
1/128	0.00034613	0.00044903	0.00048690
1/256	0.00017307	0.00022451	0.00024345

Table II. Discretization errors for the hybridized mixed Raviart-Thomas method.

Size	116	496	2048	8320	33536	134656	539648	2160640	8646656
Iterations	23	27	30	32	32	33	33	33	33
CPU secs	0.01	0.03	0.18	0.68	3.25	15.03	60.31	308.37	1325.99

Table III. Application of the multigrid algorithm to a problem on the non-convex domain of Figure 3.

is not convex.

Next, we investigate flexibility in regard to the number of smoothings m_k. Theorem 3.1 assumes that the number of smoothings increases geometrically as we decrease the refinement level k. We now repeat the first experiment, but instead of setting $m_k = 2^{J-k}$, we now fix m_k to be one for all k. Table IV indicates that the V-cycle algorithm continues to exhibit mesh independent convergence.

Finally, we give numerical results when we replace our intergrid transfer operator with two seemingly plausible intergrid transfer operators in the variable V-cycle. These operators fail, as we shall see, but they provide insight into what one should avoid when constructing a good prolongation. Consider Algorithm 3.1 with the nonnested multilevel spaces $M_k =$

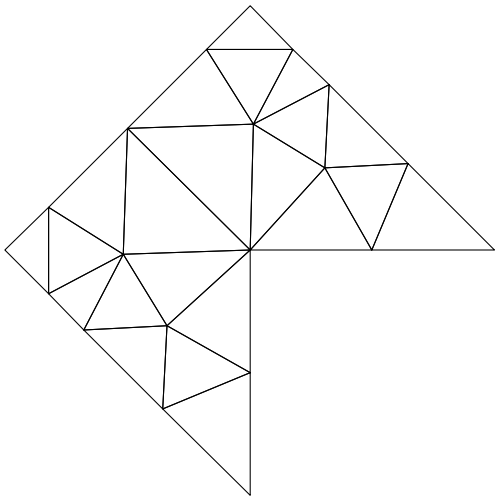

Figure 3. The non-convex domain used in experiments.

Size	74	316	1304	5296	21344	85696	343424	1374976	5502464
Iterations	21	26	31	34	34	34	35	35	35
CPU secs	0.01	0.02	0.12	0.37	1.82	8.58	34.36	134.61	533.06

Table IV. V-cycle with constant number of smoothings.

$\{\mu \in L^2(\mathscr{E}_k^i) : \mu|_e \in P_0(e), \text{ for all } e \in \mathscr{E}_k^i\}$ and non-inherited forms at every level given by $a_k(u,v) = \int_\Omega \mathcal{Q}u \cdot \mathcal{Q}v$, where now the liftings $\mathcal{Q}(\cdot)$ are defined with respect to $\mathscr{T}_k$. The forms $a_k(\cdot,\cdot)$ and a base inner product as in (9) generalized to all levels k, define the multilevel operators A_k and Q_k in the algorithm. In other words the lowest order hybridized mixed method is used to define the spaces and forms at every refinement level in the algorithm. Then we consider two different intergrid transfer operators $I_k : M_{k-1} \to M_k$, given as follows.

Consider a triangle of mesh $\mathscr{T}_{k-1}$, for instance, the triangle $T = \Delta ABC$ shown in Figure 4. Let $e = \overline{AB}$ be an interior edge of T, and $\chi^{AB} \in M_{k-1}$ be the indicator function on e. After refinement, T is divided into 4 smaller triangles which belong to mesh $\mathscr{T}_k$. Define the first prolongation $I_k^{(1)}\chi^{AB} \in M_k$ on all the finer edges as follows:

$$I_k^{(1)}\chi^{AB} = \begin{cases} \mathcal{U}(\chi^{AB}) & \text{on the 3 new edges } (\overline{DE}, \overline{DF}, \overline{EF}) \\ 1 & \text{on the 2 new edges } (\overline{AD}, \overline{BD}) \\ 0 & \text{on the other 4 new edges } (\overline{AF}, \overline{CF}, \overline{BE}, \overline{CE}) \end{cases}$$

where $\mathcal{U}(\cdot)$ is the lifting operator defined in (5), but now with respect to the mesh $\mathscr{T}_{k-1}$. The second prolongation candidate we shall consider is $I_k^{(2)} : M_{k-1} \mapsto M_k$ is defined by

$$I_k^{(2)}\chi^{AB} = \begin{cases} 1/2 & \text{on the 2 new edges } (\overline{DE}, \overline{DF}) \\ 1 & \text{on the 2 new edges } (\overline{AD}, \overline{BD}) \\ 0 & \text{on the other 5 new edges } (\overline{AF}, \overline{CF}, \overline{BE}, \overline{CE}, \overline{EF}) \end{cases}$$

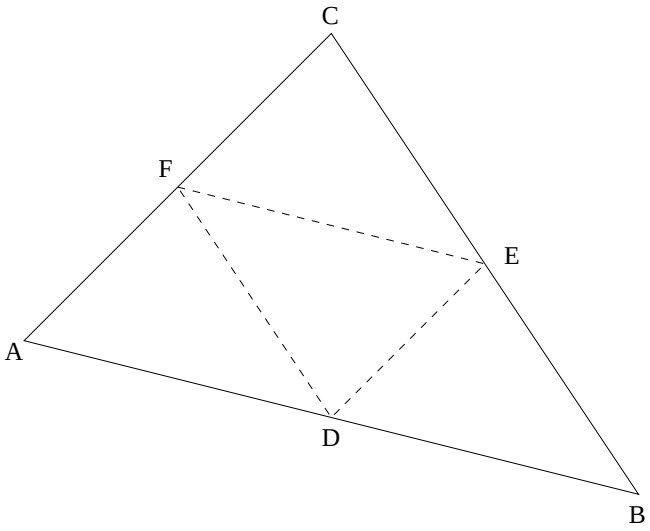

Figure 4. A refined triangle with the notations used to describe the failed intergrid transfer operators.

Size	MG iteration counts	
	with $I_k^{(1)}$	with $I_k^{(2)}$
74	34	34
316	71	81
1304	*	*
5296	*	*
21344	*	*

Table V. Failure of certain intergrid operators. (An entry * indicates that the iteration diverged.)

Each of these operators gives a different multigrid algorithm. We report on the performance of the V-cycle algorithm with these two prolongation candidates and a fixed number of smoothings $m_k = 1$ in Table V. Clearly, the results are dismal. We believe that the failure is due to the fact that prolongation operators like $I_k^{(1)}$ and $I_k^{(2)}$ increase energy upon continual transfer of a coarse grid function to increasingly finer levels. In contrast, the successful prolongation I_k that we analyzed, does not increase energy, as can be seen from Condition 4.1, which we verified for our I_k.

REFERENCES

1. 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.
2. R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp., 36 (1981), pp. 35–51.
3. D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the V-cycle*, SIAM J. Numer. Anal., 20 (1983), pp. 967–975.

4. J. H. Bramble, *Multigrid Methods*, no. 294 in Pitman research notes in mathematics series, Longman Scientific & Technical, Harlow, UK, 1993.
5. J. H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp., 58 (1992), pp. 467–488.
6. J. H. Bramble, J. E. Pasciak, J. P. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp., 57 (1991), pp. 1–21.
7. 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.
8. 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.
9. 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.
10. ———, *Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problems*, Math. Comp., 73 (2004), pp. 1041–1066 (electronic).
11. 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.
12. 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).
13. 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.
14. B. Cockburn and J. Gopalakrishnan, *A characterization of hybridized mixed methods for the Dirichlet problem*, SIAM J. Numer. Anal., 42 (2004), pp. 283–301.
15. ———, *Error analysis of variable degree mixed methods for elliptic problems via hybridization*, Math. Comp., 74 (2005), pp. 1653–1677 (electronic).
16. B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and conforming Galerkin methods for second order elliptic problems*, Submitted, (2007).
17. V. A. Dobrev, R. D. Lazarov, P. S. Vassilevski, and L. T. Zikatanov, *Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations*, Numer. Linear Algebra Appl., 13 (2006), pp. 753–770.
18. H.-Y. Duan, S.-Q. Gao, R. C. E. Tan, and S. Zhang, *A generalized BPX multigrid framework covering nonnested V-cycle methods*, Math. Comp., 76 (2007), pp. 137–152 (electronic).
19. J. Gopalakrishnan, *A Schwarz preconditioner for a hybridized mixed method*, Computational Methods in Applied Mathematics, 3 (2003), pp. 116–134.
20. P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, no. 24 in Monographs and Studies in Mathematics, Pitman Advanced Publishing Program, Marshfield, Massachusetts, 1985.
21. P.-A. Raviart and J. M. Thomas, *Primal hybrid finite element methods for* 2*nd order elliptic equations*, Math. Comp., 31 (1977), pp. 391–413.
22. 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.
23. T. Rusten and R. Winther, *Substructure preconditioners for elliptic saddle point problems*, Math. Comp., 60 (1993), pp. 23–48.
24. J. R. Shewchuk, *Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator*, in Applied Computational Geometry: Towards Geometric Engineering, M. C. Lin and D. Manocha, eds., vol. 1148 of Lecture Notes in Computer Science, Springer-Verlag, New York, May 1996, pp. 203–222. From the First ACM Workshop on Applied Computational Geometry, Philadelphia.
25. P. S. Vassilevski and J. P. Wang, *Multilevel iterative methods for mixed finite element discretizations of elliptic problems*, Numer. Math., 63 (1992), pp. 503–520.