ITERATIVE SOLVERS FOR HYBRIDIZED FINITE ELEMENT METHODS

By

SHUGUANG TAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2009

To all who nurtured my intellectual curiosity, academic interests, and sense of scholarship throughout my lifetime, making this milestone possible

ACKNOWLEDGMENTS

I am greatly thankful for enormous help from the following people who serve in my PhD supervisory committee, especially my advisor:

Prof. Jaydeep Gopalakrishnan, Department of Mathematics (advisor, committee chair)

Prof. William Hager, Department of Mathematics

Prof. Sergei Pilyugin, Department of Mathematics

Prof. Trevor Park, Department of Statistics

Prof. Alper Üngör, Department of Computer Science

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

ITERATIVE SOLVERS FOR HYBRIDIZED FINITE ELEMENT METHODS

By

Shuguang Tan

August 2009

Chair: Jaydeep Gopalakrishnan
Major: Mathematics

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. We then extend this multigrid framework to the hybridized local discontinuous Galerkin method, and yield similar mesh independent convergence results. Numerical experiments are presented to indicate the asymptotically optimal performance of our algorithm, as well as the performance comparison among different hybridized finite element methods for targeted problems.

CHAPTER 1
INTRODUCTION

The finite element method (FEM) is one of the most often used tools for numerically solving partial differential equations (PDE) on complicated domains. Depending on the importance of the properties needed in an approximate solution, one chooses the right kind of finite element method for various applications. Hybridized finite element methods have recently emerged as a powerful subclass of FEM with the ability to efficiently yield approximate solutions with interesting properties. In this study, we shall be concerned primarily with accelerating the solution process when hybridized finite elements are used.

Like other finite element methods, the hybridized finite element methods yield 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 (MG) algorithms to obtain the solution efficiently. However, the applicability of such techniques for the hybridized methods has not been investigated up to now. In this study, we shall develop multigrid algorithms for efficient solution of the systems resulting from discretization through the hybridized schemes.

When linear iterative methods are applied to solve such systems, we can think that the error consists of two parts: low frequency components and high frequency components, corresponding to small eigenvalues and large eigenvalues respectively. Classical methods such as the Gauss-Seidel iteration, reduce high frequency components quickly, but for those low frequency components, the convergence becomes really slow when the mesh size decreases. A multigrid method, on the other hand, uses a different mechanism to overcome this difficulty. Generally speaking, it maintains a sequence of grids starting from the coarsest to the finest, with the last one being the mesh on which the problem needs to be solved. A multigrid method still uses classical iterations to eliminate high frequency components on finer grids (called *smoothing*), but for low frequency components, it transfers them to coarser grids to do the reduction (called *correction*). Since the number of unknowns on coarser mesh is considerably smaller (usually only $\frac{1}{4}$ of the following finer mesh), the required amount of work is greatly reduced. That is why multigrid method has optimal performance over all other classical iterative methods.

The difficulties in adapting multigrid techniques to hybridized finite element schemes arise due to two non-standard features of hybridized methods:

1. The spaces used to find the approximate solutions consist of functions defined on the *edges* of a (two dimensional) mesh of triangles. In contrast, other finite element methods use functions defined on elements (triangles) of a mesh.
2. The approximate solution given by a hybridized method 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 make sense in a Sobolev space (without referring to a mesh).

Because of these features, when adapting multigrid algorithms to hybridized finite element methods, we must design the components of the algorithm to work with *non-nested* multilevel finite element spaces, and *non-inherited* bilinear forms. We shall use an abstract theory for the so called *variable V-cycle* algorithm which has often proved successful in analyzing multigrid adaptations to other problems with similar difficulties.

The two main ingredients of any multigrid algorithm 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, and vice versa. Since hybridized methods use functions defined on mesh edges, we must design non-trivial intergrid transfer operators. As we shall see, many of the "obvious" choices do *not* result in efficient, or even convergent, multigrid algorithms.

The next chapter is devoted to hybridized finite element methods. We give a general introduction to finite element techniques, including the "standard" FEM, and well known “mixed” FEM. Then we show the hybridized versions of the mixed method and two discontinuous Galerkin (DG) methods. In Chapter 3, we introduce multigrid theory and algorithms, and show how we adapt them to the hybridized mixed method. This is the main part of the dissertation. Then we discuss and analyze other hybridized FEMs and extend our multigrid theory and methods to them in the next two chapters. Numerical experiments are run to report the performance of the multigrid algorithms and also compare among the hybridized versions of the mixed method and the DG methods.

CHAPTER 2
HYBRIDIZED FINITE ELEMENT METHODS

The purpose of this chapter is to introduce a class of methods called hybridized finite element methods for boundary value problems, which forms the main subject of our research. To put this method in proper perspective, we begin this chapter with an introduction to the most basic and standard finite element method, which can be traced back to an early paper of Courant [21]. Then, we shall present the so-called mixed finite element method, which was introduced as a means to obtain better flux approximations [12, 31]. These discussions will provide the necessary lead in to discuss the hybridized methods in the later sections, beginning with the hybridized mixed method [2, 17], and later the hybridized discontinuous Galerkin (DG) methods [19].

Let us define some notations here first, which will be used throughout the rest of the dissertation. Let Ω be a polygonal domain in $\mathbb{R}^2$. A "subdivision" of Ω is a finite collection of closed sets called "elements" $\{K_i\}$ such that their interiors are disjoint and their union is the closure of Ω. For the finite element methods we have in mind, we require a subdivision with the further property that all elements K_i are triangles and no vertex of any triangle lies in the interior of an edge of another triangle. Such subdivisions are called "triangulations". We denote a triangulation of Ω by $\mathscr{T}_h$. The subscript h refers to

$$h = \max_{K \in \mathscr{T}_h} h_K, \qquad \text{where } h_K = \operatorname{diam}(K).$$

Let $\mathscr{E}_h$ be the set of all edges in $\mathscr{T}_h$, and $\mathscr{E}_h^i$ be the set of all interior edges in $\mathscr{T}_h$. We place the following assumption on $\mathscr{T}_h$, which is a convenient sufficient condition for the proofs of various theoretical results we shall state, but often not needed for our algorithms to make sense.

Assumption 2.1. We assume that $\mathscr{T}_h$ is *quasiuniform*, i.e., it is "shape regular" in that there is a fixed number $c_1 > 0$ such that

$$\max_{K \in \mathscr{T}_h} \frac{h_K}{\rho_K} \le c_1,$$

where ρ_K is the diameter of the largest circle inscribed in K, and furthermore, all elements are about the same size in that we additionally have another fixed number $c_2 > 0$ such that

$$c_2 h \le h_K, \quad \text{for all } K \in \mathscr{T}_h.$$

Let us now begin the discussion of the most basic finite element method.

2.1 A Standard Finite Element Method

Consider the following Dirichlet problem:

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2.1}$$

where Ω is as above, and $f \in L^2(\Omega)$. The well known weak formulation of this problem is to find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = F(v), \quad \forall v \in H_0^1(\Omega) \tag{2.2}$$

where

$$a(u,v) = \int_\Omega \vec{\nabla}\, u \cdot \vec{\nabla}\, v \, dx, \qquad F(v) = \int_\Omega f v \, dx,$$

$$H_0^1(\Omega) = \{v \in H^1(\Omega) : \quad v|_{\partial\Omega} = 0\}, \quad \text{ and}$$

$$H^1(\Omega) = \{v \in L^2(\Omega) : \quad \partial_i\, v \in L^2(\Omega), \forall i\}.$$

As usual, we have tacitly used distributional derivatives and the Sobolev trace theory in these definitions.

The essential idea of finite element methods is to replace $H_0^1(\Omega)$ with a finite-dimensional subspace V_h, and then seek an approximation u_h to the exact solution u from this subspace. For example, we can choose

$$V_h = \{v \in C(\Omega) : \quad v|_{\partial\Omega} = 0, \text{ and } v|_K \text{ is linear }, \forall K \in \mathscr{T}_h\},$$

which is obviously a subspace of $H_0^1(\Omega)$. Label the interior nodes of $\mathscr{T}_h$ by $a_1, \ldots, a_n$, and define $\phi_i \in V_h$ to be a piecewise linear function which is 1 at a_i and 0 at all other nodes. then V_h is a subspace of dimension n with $\{\phi_i\}_{i=1}^n$ as a basis.

With the basis above, we can write the approximation as

$$u_h = \sum_{i=1}^{n} c_i \phi_i$$

for some unknown coefficients $\{c_i\}$. Substituting back into (2.2), we have

$$\begin{aligned} &\int_\Omega \vec{\nabla}(\sum_{i=1}^{n} c_i \phi_i) \cdot \vec{\nabla}\, v \, dx = \int_\Omega f v \, dx, \quad \forall v \in V_h, \\ \text{i.e.} \quad &\sum_{i=1}^{n} c_i \int_\Omega \vec{\nabla}\, \phi_i \cdot \vec{\nabla}\, v \, dx = \int_\Omega f v \, dx, \quad \forall v \in V_h. \end{aligned}$$

By choosing the test function v to be one of the basis function ϕ_i, we thus obtain a matrix system

$$AC = F, \tag{2.3}$$

where $A_{i,j} = \int_\Omega \vec{\nabla}\, \phi_i \cdot \vec{\nabla}\, \phi_j \, dx$, $F_i = \int_\Omega f \phi_i \, dx$, and C is the vector of coefficients $\{c_i\}$. Since the only unknown here is C, by solving the matrix system, we can find the approximation u_h in V_h to the actual solution u. Moreover, the matrix A is symmetric and positive definite, allowing us to solve it conveniently with iterative techniques like the Conjugate Gradient method [28].

The following well known theorem [14] states the two properties most pertinent to this study, namely the conditioning of the system (2.3), and an *a priori* error estimate for the discrete approximation. We shall use $\|\cdot\|_{H^k(W)}$ and $|\cdot|_{H^k(W)}$ to denote Sobolev norms and seminorms defined on some domain W throughout this study.

Theorem 2.2. *If $\mathscr{T}_h$ satisfies Assumption 2.1, then the spectral condition number of the matrix A in (2.3) satisfies*

$$\kappa(A) \leq Ch^{-2},$$

where C is a mesh-independent constant. In addition if the exact solution u is in $H^2(\Omega)$, then

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch|u|_{H^2(\Omega)}.$$

The condition number estimate of Theorem 2.2 tell us that if we use the Conjugate Gradient method to solve (2.3), we can expect the number of iterations to grow as h decreases. Specifically, the convergence rate of the Conjugate Gradient method in the A-norm is $(\sqrt{\kappa(A)} - 1)(\sqrt{\kappa(A)} + 1)^{-1}$, so the number of iterations required to reduce the initial iterative error by some fixed tolerance factor approximately doubles when h is halved. For practically important problems, rather fine meshes are often necessary (small h), so such increasing iteration counts lead to expensive solutions strategies. Therefore, in order to efficiently solve finite systems like (2.3) we must resort either to preconditioning, or to the construction of optimal linear iterative solvers like *multigrid methods*. Our aim is to study the applicability of multigrid techniques to systems analogous to, but more complex than (2.3), arising from hybridized finite element methods (see Chapter 3).

The finite element method introduced above is an example of a Galerkin method. Since it uses a finite element space V_h of continuous functions, it is a example of a *continuous Galerkin* (CG) method. Methods that use discontinuous finite element functions will be called *discontinuous Galerkin* (DG) methods. We shall see examples of such methods in later sections. An engineering approach to the construction of finite element methods can found in [32].

2.2 Mixed Method

In some practical problems, we are also interested in the gradient of the solution, i.e., $\vec{\nabla} u$. One way to achieve this, is to apply the Galerkin method introduced above to find an approximation u_h first, and then take derivative on u_h. But this only gives us an estimation for $\vec{\nabla} u$ with low accuracy. We need a more accurate solution. Mixed method provides a better solution. It is called mixed because it seeks approximations to both u and $\vec{\nabla} u$ simultaneously. Depending on the underlying finite element spaces, there are a couple of different mixed methods. Frequently people choose the so-called Raviart-Thomas (RT) space (introduced below), which leads to the corresponding RT mixed method.

Remark 2.3. From now on, whenever referring to the mixed method (or its hybridized version), we always mean the RT mixed method (or its hybridized version). More details about hybridization will be seen later.

Still consider the model problem (2.1). If we introduce a new variable $\vec{q}$, defined as

$$\vec{q} = -\vec{\nabla}\, u,$$

then (2.1) can be reformulated as

$$\begin{aligned} \vec{q} &= -\vec{\nabla}\, u && \text{in } \Omega, \\ \nabla \cdot \vec{q} &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{2.4}$$

To distinguish, people often refer u as the *primal* solution, and $\vec{q}$ as the *flux* solution.

The RT mixed method seeks an approximation $(\vec{q}_h, u_h)$ to the exact solution $(-\vec{q}, u)$ of (2.4). Let $P_d(K)$ denote the space of polynomials on K of degree at most d, and $R_d(K)$ be the corresponding RT space, i.e.,

$$R_d(K) = P_d(K) \times P_d(K) + \vec{x} P_d(K).$$

Then the approximation $(\vec{q}_h, u_h)$ is sought in the finite element space $V_h \times W_h$ given by

$$\begin{aligned} V_h &= \{\vec{v} \in H(\mathrm{div}, \Omega) : \quad \vec{v}|_K \in R_d(K), \text{ for all } K \in \mathscr{T}_h\}, \\ W_h &= \{w \in L^2(\Omega) : \quad w|_K \in P_d(K), \text{ for all } K \in \mathscr{T}_h\}, \end{aligned}$$

and is defined by requiring that, for all $(\vec{v}, w) \in V_h \times W_h$,

$$\begin{aligned} \int_\Omega \vec{q}_h \cdot \vec{v}\, dx - \int_\Omega u_h\, \nabla \cdot \vec{v}\, dx &= 0, \\ \int_\Omega w\ \nabla \cdot \vec{q}_h\, dx &= \int_\Omega f\, w\, dx. \end{aligned} \tag{2.5}$$

Similarly as in the Galerkin method, if we choose $\{\vec{v}_i\}$ and $\{w_j\}$ to be a basis for V_h and W_h respectively, then the above weak formulation (2.5) gives rise to a matrix system of block form

$$\begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} Q \\ U \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix}, \tag{2.6}$$

where $A_{i,j} = \int_\Omega \vec{v}_i \cdot \vec{v}_j \, dx$, $B_{i,j} = -\int_\Omega w_i \vec{\nabla} \cdot \vec{v}_j \, dx$, $F_i = -\int_\Omega f \, w_i \, dx$, and Q and U are the coefficient vectors for $\vec{q}_h$ and u_h with respect to their corresponding basis $\{\vec{v}_i\}$ and $\{w_j\}$, respectively. We can then solve the linear system to get an approximation to $(\vec{q}, u)$.

Note that in contrast to (2.3), the mixed method yields the system (2.6) that is indefinite. The Conjugate Gradient method is not appropriate for such systems. We must use GMRES [33] or similar methods. While it is possible to tailor preconditioning strategies for this method, they are considerably more involved, due to the indefiniteness of the system. In contrast, the hybridized version of the mixed method, to be discussed in the next section, results in a symmetric positive definite system that gives the exactly the same solution as the mixed method. A *priori* error estimate is available for the mixed method [18, 23].

2.3 Hybridized Mixed Method

In the previous section we discussed the mixed method, which ends up with a matrix equation in (2.6). Since the system is not positive definite, solving for Q and U is not always easy. Although by elimination of Q from the equations, we can achieve a positive definite system, this will require the inversion of A, and the inverse matrix is typically a full matrix of big size. Fortunately, an advanced technique called hybridization helps us overcome this difficulty. Let us introduce the hybridized mixed method, based on the same model problem (2.4).

The method introduces an additional unknown λ_h, called *Lagrange multiplier*, which turns out to be nothing but an approximation to the trace of u on each edge $e \in \mathscr{E}_h^i$. As we will show next, its introduction can actually eliminate both $\vec{q}_h$ and u_h, reducing the system to a single matrix equation for the multiplier λ_h. In our particular case, the hybridized RT mixed (HRT) method seeks an approximation $(\vec{q}_h, u_h, \lambda_h)$ to $(\vec{q}, u, u|_{\mathscr{E}_h^i})$ in the finite element space $V_h \times W_h \times M_h$ given by

$$\begin{aligned} V_h &= \{\vec{v} \in L^2(\Omega) \times L^2(\Omega) : \quad \vec{v}|_K \in R_d(K), \quad \text{for all } K \in \mathscr{T}_h\}, \\ W_h &= \{w \in L^2(\Omega) : \quad w|_K \in P_d(K), \quad \text{for all } K \in \mathscr{T}_h\}, \\ M_h &= \{m \in L^2(\mathscr{E}_h^i) : \quad m|_e \in P_d(e), \quad \text{for all } e \in \mathscr{E}_h^i\}. \end{aligned}$$

It is defined by requiring that, for all $(\vec{v}, u, m) \in V_h \times W_h \times M_h$,

$$\begin{aligned} \int_\Omega \vec{q}_h \cdot \vec{v}\, dx - \sum_{K \in \mathscr{T}_h} \int_K u_h \nabla \cdot \vec{v}\, dx + \sum_{e \in \mathscr{E}_h^i} \int_e \lambda_h \, [\![\vec{v} \cdot \vec{n}]\!]\, ds &= 0, \\ \sum_{K \in \mathscr{T}_h} \int_K w \nabla \cdot \vec{q}_h \, dx &= \int_\Omega f\, w\, dx, \\ \sum_{e \in \mathscr{E}_h^i} \int_e m \, [\![\vec{q}_h \cdot \vec{n}]\!]\, ds &= 0. \end{aligned} \tag{2.7}$$

Here $[\![\cdot]\!]$ is called the *jump* operator, defined as below.

Definition 2.4 (Jump). On any edge e shared by two mesh triangles K^+ and K^-, the jump of a function $\vec{v}$ on e is defined by $[\![\vec{v} \cdot \vec{n}]\!] = \vec{v}|_{K^+} \cdot \vec{n}^+ + \vec{v}|_{K^-} \cdot \vec{n}^-$, and on any edge e of the boundary, $[\![\vec{v} \cdot \vec{n}]\!] = \vec{v}|_K \cdot \vec{n}$, where $\vec{n}^\pm$ denotes the outward unit normal on $\partial K^\pm$ (see Figure 2-1).

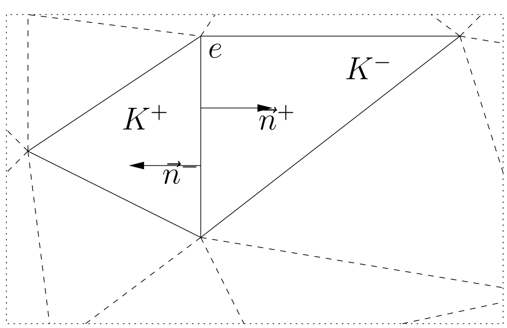

Figure 2-1. Illustration of notations

It is a well-known fact [2, 17] that the HRT method gives the exactly same solutions as the original mixed method, as stated in the following theorem.

Theorem 2.5. *The equation system (2.7) is uniquely solvable, and $(\vec{q}_h, u_h)$ in (2.7) coincide with $(\vec{q}_h, u_h)$ in (2.4) from the mixed method.*

Observing (2.7), it may seem that we need to solve an even more complicated equation system than the original mixed method. But that is really not the case. To explain why, let us define two so-called *lifting operators* as follows.

Given every $m \in L^2(\mathscr{E}_h)$, the first lifting pair of functions $(\vec{\mathfrak{Q}}^{\mathrm{RT}} m, \mathfrak{U}^{\mathrm{RT}} m) \in V_h \times W_h$ satisfies

$$\begin{aligned} \int_\Omega \vec{\mathfrak{Q}}^{\mathrm{RT}} m \cdot \vec{v}\, dx - \sum_{K\in\mathscr{T}_h} \int_K \mathfrak{U}^{\mathrm{RT}} m\, \nabla\cdot\vec{v}\, dx &= -\sum_{e\in\mathscr{E}_h}\int_e m\, [\![\vec{v}\cdot\vec{n}]\!]\, ds, \\ \sum_{K\in\mathscr{T}_h}\int_K w\,\nabla\cdot\vec{\mathfrak{Q}}^{\mathrm{RT}} m\, dx &= 0, \end{aligned} \tag{2.8}$$

and given every $f \in L^2(\Omega)$, the second lifting pair of functions $(\vec{\mathfrak{Q}}^{\mathrm{RT}} f, \mathfrak{U}^{\mathrm{RT}} f) \in V_h \times W_h$ satisfies

$$\begin{aligned} \int_\Omega \vec{\mathfrak{Q}}^{\mathrm{RT}} f \cdot \vec{v}\, dx - \sum_{K\in\mathscr{T}_h} \int_K \mathfrak{U}^{\mathrm{RT}} f\, \nabla\cdot\vec{v}\, dx &= 0, \\ \sum_{K\in\mathscr{T}_h}\int_K w\,\nabla\cdot\vec{\mathfrak{Q}}^{\mathrm{RT}} f\, dx &= \int_\Omega f\, w\, dx, \end{aligned} \tag{2.9}$$

for all $(\vec{v}, w) \in V_h \times W_h$.

One thing to know is the difference between the FEM spaces for the RT method and the HRT method. In case of the RT method, $V_h \subset H(\mathrm{div}, \Omega)$ implies that functions inside it need to have continuous normal components across boundaries, while for the HRT method, this restriction is lifted by simply requiring $V_h \subset L^2(\Omega) \times L^2(\Omega)$. Also note that these mappings defined above are uniquely determined on each element because of the surjectivity of the map $(\nabla\cdot\,) : V_h \mapsto W_h$ restricted to an element. More importantly, since the functions in $V_h \times W_h$ have *no* continuity constraints across elements, the computation of these lifting pairs can be done element by element in a decoupled way. Such computations, being *local*, are inexpensive.

In fact, on each element $K \in \mathscr{T}_h$, the lifting $(\vec{\mathcal{Q}}^{\mathrm{RT}} m, \mathcal{U}^{\mathrm{RT}} m)$ can be considered as a result of a one element discretization of the following boundary value problem

$$\begin{aligned} \vec{q}_m &= -\nabla u_m && \text{in } K, \\ \nabla \cdot \vec{q}_m &= 0 && \text{in } K, \\ u_m &= m && \text{on } \partial K, \end{aligned}$$

and the mapping $(\vec{\mathcal{Q}}^{\mathrm{RT}} f, \mathcal{U}^{\mathrm{RT}} f)$ is an approximation to the solution of

$$\begin{aligned} \vec{q}_f &= -\nabla u_f && \text{in } K, \\ \nabla \cdot \vec{q}_f &= f && \text{in } K, \\ u_f &= 0 && \text{on } \partial K. \end{aligned}$$

Once the lifting operators are defined, we have the following theorem revealing the relationship between the numerical solutions and the lifting functions [17].

Theorem 2.6 (Characterization of $\vec{q}_h, u_h$, and λ_h for HRT)**.** *Let $(\vec{q}_h, u_h, \lambda_h)$ be the solution of the HRT method (2.7). Then*

$$\vec{q}_h = \vec{\mathcal{Q}}^{\mathrm{RT}} \lambda_h + \vec{\mathcal{Q}}^{\mathrm{RT}} f \qquad \textit{and} \qquad u_h = \mathcal{U}^{\mathrm{RT}} \lambda_h + \mathcal{U}^{\mathrm{RT}} f. \tag{2.10}$$

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

$$a_h^{\mathrm{RT}}(\lambda_h, m) = b_h^{\mathrm{RT}}(m) \qquad \textit{for all}\, m \in M_h, \tag{2.11}$$

where

$$a_h^{\mathrm{RT}}(\lambda_h, m) = \int_\Omega \vec{\mathcal{Q}}^{\mathrm{RT}} \lambda_h \cdot \vec{\mathcal{Q}}^{\mathrm{RT}} m \, dx$$

and

$$b_h^{\mathrm{RT}}(m) = \int_\Omega f\, \mathcal{U}^{\mathrm{RT}} m \, dx.$$

Note that if we let $\{\phi_i\}_{i=1}^n$ be a basis for M_h, and $\lambda_h = \sum_{i=1}^n c_i\phi_i$ for some unknown coefficients $\{c_i\}$, then (2.11) results in a matrix system

$$A^{\mathrm{RT}} C = B^{\mathrm{RT}}$$

where

$$A_{i,j}^{\mathrm{RT}} = \int_\Omega \vec{\mathcal{Q}}^{\mathrm{RT}}\phi_i \cdot \vec{\mathcal{Q}}^{\mathrm{RT}}\phi_j \, dx, \qquad C_i = c_i$$

and

$$B_i^{\mathrm{RT}} = \int_\Omega f\,\mathcal{U}^{\mathrm{RT}}\phi_i \, dx$$

There are many choices for a basis in M_h, for example, if we label the interior edges in $\mathscr{E}_h^i$ by $e_1, e_2, \ldots, e_N$, then $\{p_{s,t} \in L^2(\mathscr{E}_h^i), 1 \le s \le N, 0 \le t \le k : p_{s,t} = x^t$ on e_s, and 0 elsewhere$\}$ are such a candidate.

By virtue of this theorem, instead of solving (2.7) directly, we can solve (2.11) to find the Lagrange multiplier λ_h, and then use (2.10) to recover the $\vec{q}_h$ and u_h. This recovery is *local*, as the application of $\vec{\mathcal{Q}}^{\mathrm{RT}}$ and $\mathcal{U}^{\mathrm{RT}}$ are local operations, and consequently of negligible cost, compared to the global inversion required to find λ_h.

The variational characterization (2.11) should be compared with (2.2). While (2.2) makes sense not only on the finite element space, but also on the whole Sobolev space $H^1(\Omega)$, the formulation (2.11) for the hybridized RT method only makes sense on the specific mesh under consideration. This is an important difference that we will need to keep in mind during multigrid analysis.

Note that compared to the mixed method, the hybridized version has the following advantages:

1. The size of the matrix system coming from (2.11) is smaller than the one from the mixed method (2.4) since the space M_h is defined on $\mathscr{E}_h$, not Ω.
2. The matrix from (2.11) is symmetric positive definite, and hence we can solve it with fast iterative techniques such as the Conjugate Gradient method.

3. Once λ_h is computed, the other components of the solution pair, namely $\vec{\mathcal{Q}}^{\mathrm{RT}} f$,$\vec{\mathcal{Q}}^{\mathrm{RT}} g$,$\mathcal{U}^{\mathrm{RT}} f$ and $\mathcal{U}^{\mathrm{RT}} g$, can be computed inexpensively in a completely local fashion (element by element), as seen from (2.10).
4. λ_h actually contains more information and can be used to obtain a locally post-processed solution of enhanced accuracy, as shown in [2].

2.4 Hybridized Discontinuous Galerkin Method

Recall that in the mixed method, the space V_h are chosen such that the functions in it have continuous normal components across the interior mesh faces. If we release the requirement for continuity, then the matrix A from (2.6) can be diagonalized, which makes it easier to compute the inverse. such a finite element method is called discontinuous Galerkin (DG) method. Similar as in the mixed method, by hybridizing DG methods, we can get even better results.

We consider a more general version of the model problem (2.4):

$$\begin{aligned} \vec{q} &= -\vec{\nabla} u && \text{in } \Omega, \\ \nabla \cdot \vec{q} &= f && \text{in } \Omega, \\ u &= g && \text{on } \partial\Omega, \end{aligned} \tag{2.12}$$

where $f \in L^2(\Omega)$ and $g \in H^{1/2}(\partial\Omega)$.

There are two kinds of important DG methods: local discontinuous Galerkin method and interior penalty method. Based on the general framework of the hybridized mixed method, we will briefly introduce the hybridized versions of these two methods.

Remark 2.7. We will slightly abuse our notations for simplicity of writing. From now on, HDG means the hybridized local discontinuous Galerkin method, and HIP means the hybridized interior penalty method. Recall HRT represents the hybridized Raviart-Thomas mixed method.

Before we continue, let us make a few conventions.

Notation 2.8. Throughout, constants that do not depend on h_K are generically denoted by C (with or without subscripts). Their value may differ at different occurrences, and may depend on the the shape regularity of the mesh. Any dependencies on other special variables, such as τ (defined later in § 2.4.1) will always be explicitly mentioned.

Notation 2.9. For any domain T, we write

$$(u, v)_T = \int_T uv,$$

and $\|\cdot\|_T$ is the corresponding norm. Depending on the context, we may use either form. We also define

$$\|\lambda\|_{h,T} = \left(\sum_{K \in \mathscr{T}_h,\, K \subseteq \bar{T}} \|\lambda\|^2_{L^2(\partial K)} h_K \right)^{1/2}$$

and

$$\|\!|\lambda|\!\|_{h,T} = \left(\sum_{K \in \mathscr{T}_h,\, K \subseteq \bar{T}} \|\lambda - m_K(\lambda)\|^2_{L^2(\partial K)} \frac{1}{h_K} \right)^{1/2},$$

where $m_K(\lambda) = \frac{1}{|\partial K|} \int_{\partial K} \lambda \, ds$. When the domain under consideration is Ω, we drop Ω as a subscript in notations, e.g., we use $\|\cdot\|_h$ and $\|\!|\cdot|\!\|_h$ to denote $\|\cdot\|_{h,\Omega}$ and $\|\!|\cdot|\!\|_{h,\Omega}$, respectively.

2.4.1 HDG Method

First, the finite element space $V_h \times W_h \times M_h$ for the HDG method is defined as below:

$$\begin{aligned}
V_h &= \{\vec{v} \in L^2(\Omega) \times L^2(\Omega) : \quad \vec{v}|_K \in V(K), \quad \text{for all } \ K \in \mathscr{T}_h\}, \\
W_h &= \{w \in L^2(\Omega) : \quad w|_K \in W(K), \quad \text{for all } \ K \in \mathscr{T}_h\}, \\
M_h &= \{m \in L^2(\mathscr{E}_h^i) : \quad m|_e \in P_d(e), \quad \text{for all } \ e \in \mathscr{E}_h^i\},
\end{aligned}$$

where $V(K) \times W(K)$ is one of the following choices:

$$P_d(K) \times P_d(K) \times P^{k-1}(K), \qquad k \geq 1 \text{ and } \tau_K \geq 0 \text{ on } \partial K, \tag{1}$$

$$P_d(K) \times P_d(K) \times P_d(K), \qquad k \geq 0 \text{ and } \tau_K > 0 \text{ on at least one face of } K, \tag{2}$$

$$P^{k-1}(K) \times P^{k-1}(K) \times P_d(K), \qquad k \geq 1 \text{ and } \tau_K > 0 \text{ on } \partial K, \tag{3}$$

and the so-called *stabilization parameter* τ is a nonnegative constant on each face of $\mathscr{E}_h$, allowed to be double-valued on $\mathscr{E}_h^i$ with two branches $\tau^+ = \tau_{K^+}$ and $\tau^- = \tau_{K^-}$ defined on the edge e shared by the mesh elements K^+ and K^-.

Then, similarly as in the HRT method discussed in § 2.3, we define two lifting operators locally on each mesh element K.

Given every $m \in L^2(\mathscr{E}_h)$ and every $K \in \mathscr{T}_h$, the first lifting pair of functions $(\vec{\mathfrak{Q}}^{\mathrm{DG}}m, \mathfrak{U}^{\mathrm{DG}}m) \in V_h \times W_h$ restricted on K is in $V(K) \times W(K)$, and satisfies:

$$\begin{aligned}(\vec{\mathfrak{Q}}^{\mathrm{DG}}m, \vec{v})_K - (\mathfrak{U}^{\mathrm{DG}}m, \nabla\cdot\vec{v})_K &= -(m, \vec{v}\cdot\vec{n})_{\partial K}, \quad &&\text{for all } \vec{v} \in V(K),\\ -(\vec{\nabla}\, w, \vec{\mathfrak{Q}}^{\mathrm{DG}}m)_K + (w, (\vec{\mathfrak{Q}}^{\mathrm{DG}}m\cdot\vec{n} + \tau_K(\mathfrak{U}^{\mathrm{DG}}m - m))_{\partial K} &= 0, \quad &&\text{for all } w \in W(K).\end{aligned} \tag{2.13}$$

Given every $f \in L^2(\Omega)$ and every $K \in \mathscr{T}_h$, the second lifting pair of functions $(\vec{\mathfrak{Q}}^{\mathrm{DG}}f, \mathfrak{U}^{\mathrm{DG}}f) \in V_h \times W_h$ restricted on K is in $V(K) \times W(K)$, and satisfies:

$$\begin{aligned}(\vec{\mathfrak{Q}}^{\mathrm{DG}}f, \vec{v})_K - (\mathfrak{U}^{\mathrm{DG}}f, \nabla\cdot\vec{v})_K &= 0, \quad &&\text{for all } \vec{v} \in V(K),\\ -(\vec{\nabla}\, w, \vec{\mathfrak{Q}}^{\mathrm{DG}}f)_K + (w, (\vec{\mathfrak{Q}}^{\mathrm{DG}}f\cdot\vec{n} + \tau_K(\mathfrak{U}^{\mathrm{DG}}f))_{\partial K} &= (f, w)_K, \quad &&\text{for all } w \in W(K).\end{aligned} \tag{2.14}$$

The HDG method can be viewed as seeking the approximation $(\vec{q}_h, u_h, \lambda_h)$ in $(V_h \times W_h \times M_h)$ satisfying

$$\begin{aligned}(\vec{q}_h, \vec{v})_\Omega - (u_h, \nabla\cdot\vec{v})_\Omega + \sum_{K\in\mathscr{T}_h}(\lambda_h, \vec{v}\cdot\vec{n})_{\partial K} &= -(g, \vec{v}\cdot\vec{n})_{\partial\Omega} \qquad &&\text{for all } \vec{v} \in V_h,\\ -(\vec{q}_h, \vec{\nabla}\, w)_\Omega + \sum_{K\in\mathscr{T}_h}(w, \widehat{q}_h\cdot\vec{n})_{\partial K} &= (f, w)_\Omega &&\text{for all } w \in W_h,\\ \sum_{K\in\mathscr{T}_h}(m, \widehat{q}_h\cdot\vec{n})_{\partial K} &= 0 &&\text{for all } m \in M_h,\end{aligned} \tag{2.15}$$

where $\widehat{q}_h \triangleq (\frac{\tau^-}{\tau^-+\tau^+})\vec{q}_h|_{K^+} + (\frac{\tau^+}{\tau^-+\tau^+})\vec{q}_h|_{K^-} + (\frac{\tau^+\tau^-}{\tau^-+\tau^+})\,[\![u_h\vec{n}]\!]$ (see Definition 2.4 for the jump operator $[\![\cdot]\!]$). A similar characterization theorem as for the HRT method is given as follows, with additional terms involving g:

Theorem 2.10 (Characterization of $\vec{q}_h, u_h$, and λ_h for HDG)**.** *Let $(\vec{q}_h, u_h, \lambda_h)$ be the solution of the HDG method (2.15). Then*

$$\vec{q}_h = \vec{\mathcal{Q}}^{\text{DG}}\lambda_h + \vec{\mathcal{Q}}^{\text{DG}}f + \vec{\mathcal{Q}}^{\text{DG}}g \qquad \textit{and} \qquad u_h = \mathcal{U}^{\text{DG}}\lambda_h + \mathcal{U}^{\text{DG}}f + \mathcal{U}^{\text{DG}}g. \tag{2.16}$$

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

$$a_h^{\text{DG}}(\lambda_h, m) = b_h^{\text{DG}}(m) \quad \textit{for all } m \in M_h, \tag{2.17}$$

where

$$a_h^{\text{DG}}(\lambda_h, m) = (\vec{\mathcal{Q}}^{\text{DG}}\lambda_h, \vec{\mathcal{Q}}^{\text{DG}}m)_\Omega + (1, [\![(\mathcal{U}^{\text{DG}}\lambda_h - \lambda_h)\tau(\mathcal{U}^{\text{DG}}m - m)\vec{n}]\!]_{\mathscr{E}_h},$$

and

$$b_h^{\text{DG}}(m) = (g, \vec{\mathcal{Q}}^{\text{DG}}m \cdot \vec{n})_{\partial\Omega} + (g, \tau\mathcal{U}^{\text{DG}}m)_{\partial\Omega} + (f, \mathcal{U}^{\text{DG}}m)_\Omega.$$

By choosing an appropriate basis $\{\phi_i\}$ of M_h such that $\lambda_h = \sum_{i=1}^n c_i\phi_i$ for some unknown coefficients $\{c_i\}$, (2.17) results in a matrix system

$$A^{\text{DG}}C = B^{\text{DG}}$$

where

$$A_{i,j}^{\text{DG}} = (\vec{\mathcal{Q}}^{\text{DG}}\phi_i, \vec{\mathcal{Q}}^{\text{DG}}\phi_j)_\Omega + (1, [\![(\mathcal{U}^{\text{DG}}\phi_i - \phi_i)\tau(\mathcal{U}^{\text{DG}}\phi_j - \phi_j)\vec{n}]\!])_{\mathscr{E}_h},$$

and

$$B_i^{\text{DG}} = (g, \vec{\mathcal{Q}}^{\text{DG}}\phi_i \cdot \vec{n})_{\partial\Omega} + (g, \tau\mathcal{U}^{\text{DG}}m)_{\partial\Omega} + (f, \mathcal{U}^{\text{DG}}m)_\Omega.$$

More details about the HDG method will be discussed in later chapters.

2.4.2 HIP Method

First, the finite element space $V_h \times W_h \times M_h$ for the HIP method is defined as below:

$$\begin{aligned}
V_h &= \{\vec{v} \in L^2(\Omega) \times L^2(\Omega) : \quad \vec{v}|_K \in P_d(K) \times P_d(K), \quad \text{for all } \ K \in \mathscr{T}_h\},\\
W_h &= \{w \in L^2(\Omega) : \quad w|_K \in P_d(K), \quad \text{for all } \ K \in \mathscr{T}_h\},\\
M_h &= \{m \in L^2(\mathscr{E}_h^i) : \quad m|_e \in P_d(e), \quad \text{for all } \ e \in \mathscr{E}_h^i\}.
\end{aligned}$$

Then, similarly as in the HRT method discussed in § 2.3, we define two lifting operators locally on each mesh element K.

Given every $m \in L^2(\mathscr{E}_h)$ and every $K \in \mathscr{T}_h$, the first lifting pair of functions $(\vec{\mathcal{Q}}^{\mathrm{IP}} m, \mathcal{U}^{\mathrm{IP}} m) \in V_h \times W_h$ restricted on K is in $V(K) \times W(K)$, and satisfies:

$$\begin{aligned}
(\vec{\mathcal{Q}}^{\mathrm{IP}} m, \vec{v})_K - (\mathcal{U}^{\mathrm{IP}} m, \nabla \cdot \vec{v})_K &= -(m, \vec{v} \cdot \vec{n})_{\partial K}, \quad \text{for all } \vec{v} \in V(K), \\
-(\vec{\nabla}\, w, \vec{\mathcal{Q}}^{\mathrm{IP}} m)_K + (w, -a\, \vec{\nabla}\, \mathcal{U}^{\mathrm{IP}} m \cdot \vec{n} + \tau_K(\mathcal{U}^{\mathrm{IP}} m - m))_{\partial K} &= 0, \quad \text{for all } w \in W(K).
\end{aligned} \tag{2.18}$$

Given every $f \in L^2(\Omega)$ and every $K \in \mathscr{T}_h$, the second lifting pair of functions $(\vec{\mathcal{Q}}^{\mathrm{IP}} f, \mathcal{U}^{\mathrm{IP}} f) \in V_h \times W_h$ restricted on K is in $V(K) \times W(K)$, and satisfies:

$$\begin{aligned}
(\vec{\mathcal{Q}}^{\mathrm{IP}} f, \vec{v})_K - (\mathcal{U}^{\mathrm{IP}} f, \nabla \cdot \vec{v})_K &= 0, \quad \text{for all } \vec{v} \in V(K), \\
-(\vec{\nabla}\, w, \vec{\mathcal{Q}}^{\mathrm{IP}} f)_K + (w, -a\, \vec{\nabla}\, \mathcal{U}^{\mathrm{IP}} f \cdot \vec{n} + \tau_K \mathcal{U}^{\mathrm{IP}} f)_{\partial K} &= (f, w)_K, \quad \text{for all } w \in W(K).
\end{aligned} \tag{2.19}$$

As in the HDG method, τ is double-valued on each interior mesh face of $\mathscr{E}_h$.

The HIP method can be viewed as seeking the approximation $(\vec{q}_h, u_h, \lambda_h)$ in $(V_h \times W_h \times M_h)$ satisfying

$$\begin{aligned}
(\vec{q}_h, \vec{v})_\Omega - (u_h, \nabla \cdot \vec{v})_\Omega + \sum_{K \in \mathscr{T}_h} (\lambda_h, \vec{v} \cdot \vec{n})_{\partial K} &= -(g, \vec{v} \cdot \vec{n})_{\partial\Omega} \qquad \text{for all } \vec{v} \in V_h, \\
-(\vec{q}_h, \vec{\nabla}\, w)_\Omega + \sum_{K \in \mathscr{T}_h} (w, \widehat{q}_h \cdot \vec{n})_{\partial K} &= (f, w)_\Omega \qquad \text{for all } w \in W_h, \\
\sum_{K \in \mathscr{T}_h} (m, \widehat{q}_h \cdot \vec{n})_{\partial K} &= 0 \qquad \text{for all } m \in M_h,
\end{aligned} \tag{2.20}$$

where $\widehat{q}_h \triangleq -(\frac{\tau^-}{\tau^- + \tau^+}) a^+ \vec{\nabla}\, u_h|_{K+} - (\frac{\tau^+}{\tau^- + \tau^+}) a^- \vec{\nabla}\, u_h|_{K-} + (\frac{\tau^+ \tau^-}{\tau^- + \tau^+}) [\![u_h \vec{n}]\!]$. A similar characterization theorem as (2.6) is given as follows:

Theorem 2.11 (Characterization of $\vec{q}_h, u_h$, and λ_h for HIP)**.** *Let $(\vec{q}_h, u_h, \lambda_h)$ be the solution of the HIP method (2.20). Then*

$$\vec{q}_h = \vec{\mathcal{Q}}^{\mathrm{IP}} \lambda_h + \vec{\mathcal{Q}}^{\mathrm{IP}} f + \vec{\mathcal{Q}}^{\mathrm{IP}} g \qquad \textit{and} \qquad u_h = \mathcal{U}^{\mathrm{IP}} \lambda_h + \mathcal{U}^{\mathrm{IP}} f + \mathcal{U}^{\mathrm{IP}} g. \tag{2.21}$$

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

$$a_h^{\mathrm{IP}}(\lambda_h, m) = b_h^{\mathrm{IP}}(m) \quad \text{for all } m \in M_h, \tag{2.22}$$

where

$$a_h^{\mathrm{IP}}(\lambda_h, m) = (\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}\lambda_h, \vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}m)_{\mathscr{T}_h} - (1, [\![((\mathcal{U}^{\mathrm{IP}}\lambda_h - \lambda_h)\,\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}m + (\mathcal{U}^{\mathrm{IP}}m - m)\,\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}\lambda_h)\vec{n}]\!])_{\mathscr{E}_h}$$
$$+ (1, [\![(\mathcal{U}^{\mathrm{IP}}\lambda_h - \lambda_h)\tau(\mathcal{U}^{\mathrm{IP}}m - m)\vec{n}]\!])_{\mathscr{E}_h},$$

and

$$b_h^{\mathrm{IP}}(m) = (f, \mathcal{U}^{\mathrm{IP}}m)_\Omega + (g, -\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}m \cdot \vec{n} + \tau\mathcal{U}^{\mathrm{IP}}m)_{\partial\Omega}.$$

By choosing an appropriate basis $\{\phi_i\}$ of M_h such that $\lambda_h = \sum_{i=1}^n c_i\phi_i$ for some unknown coefficients $\{c_i\}$, (2.20) results in a matrix system

$$A^{\mathrm{IP}}C = B^{\mathrm{IP}}$$

where

$$A_{i,j}^{\mathrm{IP}} = (\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}\phi_i, \mathcal{U}^{\mathrm{IP}}\phi_j)_{\mathscr{T}_h} - (1, [\![((\mathcal{U}^{\mathrm{IP}}\phi_i - \phi_i)\,\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}\phi_j + (\mathcal{U}^{\mathrm{IP}}\phi_j - \phi_j)\,\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}\phi_i)\vec{n}]\!])_{\mathscr{E}_h}$$
$$+ (1, [\![(\mathcal{U}^{\mathrm{IP}}\phi_i - \phi_i)\tau(\mathcal{U}^{\mathrm{IP}}\phi_j - \phi_j)\vec{n}]\!])_{\mathscr{E}_h},$$

and

$$B_i^{\mathrm{IP}} = (f, \mathcal{U}^{\mathrm{IP}}\phi_i)_\Omega + (g, -\vec{\nabla}\,\mathcal{U}^{\mathrm{IP}}\phi_i \cdot \vec{n} + \tau\mathcal{U}^{\mathrm{IP}}\phi_i)_{\partial\Omega}.$$

More details regarding the HIP method will be discussed later.

CHAPTER 3
MULTIGRID ALGORITHM FOR THE HRT METHOD

In this chapter, we will develop a multigrid algorithm for solving the linear system arising from the hybridized mixed method. As we have already seen in Section 2.3, all solution components can be recovered once we find λ_h in M_h satisfying

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

However, just like in the continuous Galerkin method, the stiffness matrix of this mesh dependent variational system has condition number that grows like $O(h^{-2})$, as stated in the following theorem. Its proof can be found in [24].

Theorem 3.1. *Suppose Assumption 2.1 holds. Then the spectral condition number of* A^{RT} *satisfies*

$$\kappa(A^{\text{RT}}) \leq Ch^{-2}.$$

This growth in condition number implies that many classical iterative schemes, such as the Jacobi method [1], the Gauss-Seidel method [27], or the un-preconditioned Conjugate Gradient method [28] will yield increasing iteration counts as h goes to 0. Therefore, we need iterative solution strategies that do not deteriorate in performance when the mesh size decreases. In this section, we give one such technique. The algorithm fits into the abstract framework of [9, 10] 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.

The results of this chapter have appeared in [25].

3.1 Multigrid 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 in each mesh element (triangle) 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

$$M_k = \{v \in C(\Omega) : \quad v|_{\partial\Omega} = 0,\ v|_K \in P_1(K) \text{ for all } K \in \mathscr{T}_{k+1}\},$$

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

$$M_J = \{\mu \in L^2(\mathscr{E}_J) : \quad \mu|_e \in P_d(e) \text{ for all } e \in \mathscr{E}_J\}$$

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

$$(\eta, \mu)_J = \sum_{K \in \mathscr{T}_J} \frac{|K|}{|\partial K|} \int_{\partial K} \eta\, \mu, \tag{3.1}$$

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

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

where $\vec{\mathbb{Q}}^{\mathrm{RT}}$ 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

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

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* $I_k : M_{k-1} \to M_k$, $(k = 2, \dots, J)$ by

$$I_k v = \begin{cases} v, & \text{for } k < J, \\ v|_{\mathscr{E}_J}, & \text{for } k = J \text{ and } d > 0, \\ \Pi_{M_J}(v|_{\mathscr{E}_J}), & \text{for } k = J \text{ and } d = 0, \end{cases} \tag{3.2}$$

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 3.3, 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 3.3) as we proceed to the coarser levels.

Algorithm 3.2 (Variable V-cycle for the HRT method). 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 $l = m_k + 1, \dots, 2m_k$,

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

(f) Finally, define the next iterate by setting

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

This algorithm can be used as either a linear iteration or 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 [10] that $\mathcal{E}_J$ is a linear operator admitting a recursive expression. Using the abstract theory of [9, 10], we prove that this iterative error decreases geometrically at a mesh-independent rate, as stated in the next theorem.

Theorem 3.3. *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 3.2 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 dissertation. 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 also briefly explain how to use Algorithm 3.2 as a preconditioner. 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 an iterative operator A_J, e.g., the 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 3.3 shows that $\kappa(B_J A_J)$ is bounded above and below by mesh independent constants. Hence B_J is an optimal preconditioner.

3.2 Proof of the Convergence Result

This section is devoted to the proof of Theorem 3.3. We shall use the abstract multigrid theory of [9, 10] 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 3.3 is complete.

Condition 3.4 (Prolongation norm). For all $k = 1, \dots, J$,

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

Condition 3.5 (Regularity & Approximation). There exist $0 < \alpha \le 1$ and $C > 0$ such that

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

where λ_k is the eigenvalue of A_k with maximal norm, and $P_{k-1} : M_k \to M_{k-1}$ 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 3.6 (Smoothing). There exists $\omega > 0$ such that

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

where $\tilde{R}_k = R_k + R_k^t - R_k A_k R_k^t$.

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

3.2.1 Verification of Condition 3.4

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

$$\vec{\mathfrak{Q}}^{\mathrm{RT}}(I_J v_{J-1}) = -\vec{\nabla}\, v_{J-1}. \tag{3.3}$$

Proof. Let the divergence free subspace of the RT space be denoted by

$$R_d^0(K) = \{\vec{q} \in R_d(K) : \quad \nabla \cdot (\vec{q}|_K) = 0\} \tag{3.4}$$

for any $K \in \mathscr{T}_J$. Then, by the definition of the lifting $\vec{\mathfrak{Q}}^{\mathrm{RT}}(\cdot)$,

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

First consider the case $d = 0$. Then, by (3.2), the right hand side above can be rewritten as

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

where the last two equalities follow because $\vec{r} \cdot \vec{n}$ is piecewise constant when $d = 0$, and by integration by parts, respectively. Since both $\vec{\mathfrak{Q}}^{\mathrm{RT}}(I_J v_{J-1})|_K$ and $\vec{\nabla}\, v_{J-1}|_K$ 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 (3.2), we have

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

so the proof can be completed as before.

The verification of Condition 3.4 is now completed by observing that because of Lemma 3.7,

$$a_J(I_J v, I_J v) = (\vec{\mathfrak{Q}}^{\mathrm{RT}}(I_J v), \vec{\mathfrak{Q}}^{\mathrm{RT}}(I_J v)) = (\vec{\nabla}\, v_{J-1}, \vec{\nabla}\, v_{J-1}) = a_{J-1}(v_{J-1}, v_{J-1}).$$

Here $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product and is not to be confused with $(\cdot, \cdot)_J$ defined in (3.1).

3.2.2 Verification of Condition 3.5

Inequalities like that of Condition 3.5 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 3.5 holds for $k = 0, 1 \ldots, J-1$ [6], which is part of the standard full regularity based proofs of multigrid convergence for the continuous Galerkin method [3, 5, 10]. So we only need to verify Condition 3.5 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 $\mathfrak{U}^{\mathrm{RT}}(\cdot)$ defined earlier.

Lemma 3.8. *For all $w \in P_1(K)$, we have*

$$\mathfrak{U}^{\mathrm{RT}}(I_J w) = \Pi_{W_h} w$$

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

Proof. Given $w \in P_1(K)$, by the definition of the lifting operators in (2.8), we have

$$\int_K \vec{\mathfrak{Q}}^{\mathrm{RT}}(I_J w) \cdot \vec{r} - \int_K \mathfrak{U}^{\mathrm{RT}}(I_J w)\; \nabla \cdot \vec{r} = -\int_{\partial K} I_J w\; (\vec{r} \cdot \vec{n}), \qquad \forall \vec{r} \in R_d(K).$$

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 $\vec{r} \cdot \vec{n}$ takes a constant value on each edge. Therefore, using integration by parts, we obtain

$$\int_K \vec{\mathcal{Q}}^{\mathrm{RT}}(I_J w) \cdot \vec{r} - \int_K \mathcal{U}^{\mathrm{RT}}(I_J w)\, \nabla \cdot \vec{r} = -\int_K \vec{\nabla}\, w \cdot \vec{r} - \int_K w\ \nabla \cdot \vec{r}$$

Since $\vec{\mathcal{Q}}^{\mathrm{RT}}(I_J w) = -\vec{\nabla}\, w$ by Lemma 3.7, this implies

$$\int_K (\mathcal{U}^{\mathrm{RT}}(I_J w) - w)\ \nabla \cdot \vec{r} = 0, \qquad \forall \vec{r} \in R_d(K).$$

The lemma now follows, since $\nabla \cdot\ : R_d(K) \to P_d(K)$ 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

$$\begin{aligned} \int_K \psi(p,\lambda)\, s &= \int_K p\, s, \quad \forall s \in P_d(K), \text{ and} \\ \int_e \psi(p,\lambda)\, \mu &= \int_e \lambda \mu, \quad \forall \mu \in P_{d_+-1}(e), \end{aligned} \tag{3.5}$$

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 $\mathrm{diam}(K)/\rho_K$ over all elements K, where ρ_K denotes the diameter of the largest ball inscribed in K.

Lemma 3.9. *There is a unique $\psi(p, \lambda)$ in $L_d(K)$ satisfying (3.5). 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 (3.5) forms a square system for $\psi(p, \lambda)$. Indeed, the number of equations in the system (3.5) equals

$$\dim(P_d(K)) + 3\dim(P_{d_+-1}(e)) = \frac{1}{2}(d+1)(d+2) + 3d_+. \tag{3.6}$$

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 (3.6). Thus (3.5) is a square system.

To prove that there is a unique $\psi(p, \lambda)$ satisfying (3.5), it now suffices to show that if p and λ vanish, the only solution of (3.5) is trivial. To this end, consider a ψ in $L_d(K)$ satisfying

$$\int_K \psi s = 0, \quad \forall s \in P_d(K) \tag{3.7}$$

$$\int_e \psi \mu = 0, \quad \forall e, \forall \mu \in P_{d_+-1}(e). \tag{3.8}$$

The last equation (3.8) 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 (3.7) 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 (3.5).

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 $\hat{K}$ we have

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

due to a well-known fact that norms on a finite dimensional space are equivalent to each other. 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

$$S\lambda = \psi(\mathcal{U}^{\mathrm{RT}}\lambda, \lambda). \tag{3.9}$$

Note that $S\lambda$ is in $H(\mathrm{div}, \Omega)$ since its normal components are continuous because of (3.5). 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 (3.1).

Lemma 3.10. *For any λ in M_J, the following statements hold:*

$$C_1 \|\lambda\|_a^2 \ge \sum_{K \in \mathscr{T}_J} \left| S\lambda \right|^2_{H^1(K)} \tag{3.10}$$

$$C_2 \|\lambda\|_J \le \|S\lambda\|_{L^2(\Omega)} \le C_3 \|\lambda\|_J \tag{3.11}$$

$$\mathcal{U}^{\mathrm{RT}}\lambda = \Pi_{W_h}(S\lambda). \tag{3.12}$$

Here C_i's are mesh independent constants.

Proof. To prove (3.10), 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 3.8). Hence, for any λ, we have

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

Therefore,

$$\begin{aligned}
\| \vec{\nabla}(S\lambda)\|_{L^2(K)} &= \| \vec{\nabla} S\lambda - \vec{\nabla} S\bar{\lambda}\|_{L^2(K)}, \\
&\le Ch^{-1}\|S(\lambda-\bar{\lambda})\|_{L^2(K)} && \text{by inverse inequality} \\
&\le Ch^{-1}(\|\mathfrak{U}^{\text{RT}}(\lambda-\bar{\lambda})\|_{L^2(K)} + |\partial K|^{1/2}\|\lambda-\bar{\lambda}\|_{L^2(\partial K)}) && \text{by Lemma 3.9} \\
&\le Ch^{-1}(|\partial K|^{1/2}\|\lambda-\bar{\lambda}\|_{L^2(\partial K)}) && \text{by [18, Lemma 3.3]} \\
&\le C|\partial K|^{-1/2}\|\lambda-\bar{\lambda}\|_{L^2(\partial K)}.
\end{aligned}$$

Summing over all elements and using a norm equivalence proved in [24, Theorem 2.2], we get

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

which proves (3.10).

The proof of (3.11) is a straightforward consequence of Lemma 3.9.

The identity (3.12) is obvious from (3.5) and (3.9).

We need one more intermediate map before we can give our proof of Condition 3.5. To describe this map, first we define a ϕ_λ in M_J for every λ in M_J by

$$(S\phi_\lambda, S\mu) = (\vec{\mathfrak{Q}}^{\text{RT}}\lambda, \vec{\mathfrak{Q}}^{\text{RT}}\mu), \quad \forall \mu \in M_J. \tag{3.13}$$

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 (3.11) of Lemma 3.10. Next, let $f_\lambda = \mathfrak{U}^{\text{RT}}\phi_\lambda$. 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

$$a(\tilde{\lambda}, \mu) = (f_\lambda, \mathcal{U}^{\mathrm{RT}}\mu), \quad \forall \mu \in M_J. \tag{3.14}$$

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

Lemma 3.11. *Let $\lambda, \tilde{\lambda}$ and ϕ_λ be defined as above. Then*

$$\|S\phi_\lambda\|_{L^2(\Omega)} \leq C\|A_J\lambda\|_J \tag{3.15}$$

$$\|\lambda - \tilde{\lambda}\|_a \leq Ch_J\|A_J\lambda\|_J. \tag{3.16}$$

Proof. To prove (3.15),

$$\begin{aligned}
\|S\phi_\lambda\|_{L^2(\Omega)} &= \sup_{\mu \in M_J} \frac{(S\phi_\lambda, S\mu)}{\|S\mu\|_{L^2(\Omega)}} = \sup_{\mu \in M_J} \frac{a_J(\lambda, \mu)}{\|S\mu\|_{L^2(\Omega)}} && \text{by (3.13)} \\
&= \sup_{\mu \in M_J} \frac{(A_J\lambda, \mu)_J}{\|S\mu\|_{L^2(\Omega)}} \leq C \sup_{\mu \in M_J} \frac{(A_J\lambda, \mu)_J}{\|\mu\|_J} && \text{by (3.11)} \\
&\leq C\|A_J\lambda\|_J.
\end{aligned}$$

To prove (3.16), note that

$$a_J(\lambda, \mu) = (S\phi_\lambda, S\mu)$$

$$a_J(\tilde{\lambda}, \mu) = (\Pi_{W_h} S\phi_\lambda, S\mu),$$

where the last identity follows because in (3.14), the right hand side function is

$$f_\lambda = \mathcal{U}^{\mathrm{RT}}\phi_\lambda = \Pi_{W_h}(S\phi_\lambda) \tag{3.17}$$

by (3.12). Subtracting, and setting $\mu = \lambda - \tilde{\lambda}$, we get

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

Using the Friedrichs estimate [11] $\|u - \Pi_{W_h} u\|_{L^2(K)} \leq Ch_J |u|_{H^1(K)}$, we get

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

by an inverse inequality and (3.15). Finally, applying the estimate (3.10) of Lemma 3.10 on the last term above, and canceling the common factor, we obtain (3.16).

With these preparations, we can now finish the verification of Condition 3.5.

Proof of Condition 3.5 with $\alpha = 1$. Let λ be in M_J. We need to estimate the quantity

$$
a_J((I - I_J P_{J-1})\lambda, \lambda) = \int_\Omega \vec{\mathfrak{Q}}^{\mathrm{RT}}(\lambda - I_J P_{J-1}\lambda) \cdot \vec{\mathfrak{Q}}^{\mathrm{RT}} \lambda
$$

To estimate this, it will be useful to make the following preliminary observations:

First, by Lemma 3.7,

$$
\vec{\mathfrak{Q}}^{\mathrm{RT}}(\lambda - I_J P_{J-1}\lambda) = \vec{\mathfrak{Q}}^{\mathrm{RT}}(\lambda) + \vec{\nabla}(P_{J-1}\lambda). \tag{3.18}
$$

Second, the expression on the right hand side above satisfies

$$
(\vec{\mathfrak{Q}}^{\mathrm{RT}}\lambda + \vec{\nabla}(P_{J-1}\lambda), \vec{\nabla}\, v_{J-1})_\Omega = 0, \qquad \forall v_{J-1} \in M_{J-1}, \tag{3.19}
$$

because

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

Thus,

$$a_J((I - I_J P_{J-1})\lambda, \lambda) = \|\vec{\mathfrak{Q}}^{\mathrm{RT}}\lambda + \vec{\nabla}(P_{J-1}\lambda)\|^2, \tag{3.20}$$

and it suffices to estimate $\vec{\mathfrak{Q}}^{\mathrm{RT}}\lambda + \vec{\nabla}(P_{J-1}\lambda)$. 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 $\vec{\mathfrak{Q}}^{\mathrm{RT}}\lambda + \vec{\nabla}(P_{J-1}\lambda)$ into many terms, labeling each one as follows:

$$\begin{aligned}
\vec{\mathfrak{Q}}^{\mathrm{RT}}\lambda + \vec{\nabla}(P_{J-1}\lambda) = {} & \vec{\mathfrak{Q}}^{\mathrm{RT}}(\lambda - \tilde{\lambda}) && \cdots\cdots\cdots(\text{term A}) \\
& + \vec{\mathfrak{Q}}^{\mathrm{RT}}\tilde{\lambda} - \vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) && \cdots\cdots\cdots(\text{term B}) \\
& + \vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) + \vec{\nabla}\,\tilde{u} && \cdots\cdots\cdots(\text{term C}) \\
& + \vec{\nabla}\,P_{J-1}\tilde{\lambda} - \vec{\nabla}\,\tilde{u} && \cdots\cdots\cdots(\text{term D}) \\
& + \vec{\nabla}\,P_{J-1}(\lambda - \tilde{\lambda}), && \cdots\cdots\cdots(\text{term E})
\end{aligned} \tag{3.21}$$

where $\tilde{\lambda}$ is as defined in (3.14) and $\tilde{u}$ is the unique function in $H_0^1(\Omega)$ that solves

$$(\vec{\nabla}\,\tilde{u}, \vec{\nabla}\,v) = (f_\lambda, v), \quad \forall v \in H_0^1(\Omega).$$

Note that since we have assumed a convex domain,

$$\|\tilde{u}\|_{H^2(\Omega)} \le C\|f_\lambda\|_{L^2(\Omega)}. \tag{3.22}$$

by a well known regularity theorem [26].

The first term can be estimated by

$$\|(\text{term A})\| = \|\vec{\mathfrak{Q}}^{\mathrm{RT}}(\lambda - \tilde{\lambda})\| = \|\lambda - \tilde{\lambda}\|_a \le Ch_J\|A_J\lambda\|_J$$

by the inequality (3.16) of Lemma 3.11. For the next term, first observe that due to the characterization of Lagrange multipliers given by Theorem 2.6, $\tilde{\lambda}$ is the hybridized mixed method approximation to $\tilde{u}$. Hence by a previously established Lagrange multiplier error estimate [18, Theorem 3.1],

$$\|\vec{\mathfrak{Q}}^{\mathrm{RT}}\tilde{\lambda} - \vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u})\| \le \|\vec{q} - \Pi_R\,\vec{q}\| \tag{3.23}$$

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

$$\begin{aligned}
\|(\text{term B})\| = \|\vec{\mathfrak{Q}}^{\mathrm{RT}}\tilde{\lambda} - \vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u})\| &\le Ch_J|\tilde{u}|_{H^2(\Omega)} \\
&\le Ch_J\|f_\lambda\|_{L^2(\Omega)} \qquad \text{by (3.22)} \\
&= Ch_J\|\Pi_{W_h}(S\phi_\lambda)\|_{L^2(\Omega)} \le Ch_J\|S\phi_\lambda\|_{L^2(\Omega)} \\
&\le Ch_J\|A_J\lambda\|_J,
\end{aligned}$$

using the estimate (3.15) of Lemma 3.11.

We proceed to analyze the next term. For this, recall the divergence free subspace $R_d^0(K)$ defined in (3.4). By the definition of $\vec{\mathfrak{Q}}^{\mathrm{RT}}(\cdot)$,

$$\begin{aligned}
\int_K \vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u})\cdot\vec{r} = -\int_{\partial K}(\Pi_{M_J}\tilde{u})\vec{r}\cdot\vec{n} &= -\int_{\partial K}\tilde{u}\,\vec{r}\cdot\vec{n} \\
= -\int_K \vec{\nabla}\,\tilde{u}\cdot\vec{r}&
\end{aligned} \tag{3.24}$$

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

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

Then, on any mesh element K, its flux $-\vec{\nabla}\,\tilde{u}_{J-1}|_K$ is constant, and therefore in $R_d^0(K)$. Putting $\vec{r} = \vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) + \vec{\nabla}\,\tilde{u}_{J-1}|_K$ in (3.24), we have

$$\left(\vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) + \vec{\nabla}\,\tilde{u},\ \vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) + \vec{\nabla}\,u_{J-1}\right)_K = 0,$$

or in other words,

$$\|\vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) + \vec{\nabla}\,\tilde{u}\|_{L^2(K)}^2 = \left(\vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) + \vec{\nabla}\,\tilde{u}, \vec{\nabla}\,\tilde{u} - \vec{\nabla}\,\tilde{u}_{J-1}\right)_K.$$

By Cauchy-Schwarz inequality,

$$\begin{aligned}\|(\text{term C})\| &= \|\vec{\mathfrak{Q}}^{\mathrm{RT}}(\Pi_{M_J}\tilde{u}) + \vec{\nabla}\,\tilde{u}\|_{L^2(\Omega)} \\ &\le \|\,\vec{\nabla}(\tilde{u} - \tilde{u}_{J-1})\|_{L^2(\Omega)} \le Ch_J|\tilde{u}|_{H^2(\Omega)} \le Ch_J\|f_\lambda\|_{L^2(\Omega)} \\ &\le Ch_J\|A_J\lambda\|_J,\end{aligned} \tag{3.25}$$

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 (3.22), and (3.15) of Lemma 3.11.

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

$$\begin{aligned}(\vec{\nabla}\,P_{J-1}\tilde{\lambda}, \vec{\nabla}\,w_{J-1}) &= -(\vec{\mathfrak{Q}}^{\mathrm{RT}}\tilde{\lambda}, \vec{\nabla}\,w_{J-1}) && \text{by (3.19)} \\ &= (\vec{\mathfrak{Q}}^{\mathrm{RT}}\tilde{\lambda}, \vec{\mathfrak{Q}}^{\mathrm{RT}}(I_J w_{J-1})) && \text{by Lemma 3.7} \\ &= (f_\lambda, \mathfrak{U}^{\mathrm{RT}}(I_J w_{J-1})) && \text{by (3.14) and Theorem 2.6} \\ &= (f_\lambda, w_{J-1}) && \text{by Lemma 3.8 and (3.17).}\end{aligned}$$

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

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

by the same arguments as in (3.25).

For the final term, we first note that if we choose $\lambda = \mu$ and $v_{J-1} = P_{J-1}\mu$ in (3.19), then we have

$$\|\,\vec{\nabla}\,P_{J-1}\mu\|^2 = -(\vec{\mathfrak{Q}}^{\mathrm{RT}}\mu, \vec{\nabla}\,P_{J-1}\mu),$$

and hence, by Cauchy-Schwarz inequality,

$$\|\,\vec{\nabla}\,P_{J-1}\mu\| \le \|\vec{\mathfrak{Q}}^{\mathrm{RT}}\mu\|, \quad \forall \mu \in M_{J-1}$$

Therefore

$$\|(\text{term E})\| = \| \vec{\nabla}\, P_{J-1}(\lambda - \tilde{\lambda})\| \leq \|\vec{\mathfrak{Q}}^{\text{RT}}(\lambda - \tilde{\lambda})\| \leq Ch_J \|A_J\lambda\|_J$$

by the estimate (3.16) of Lemma 3.11.

Returning to (3.20) and combining the estimates for each of the terms above, we obtain

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

By [24, Theorem 2.3], we know that $\lambda_J \leq Ch_J^{-2}$. Hence the above inequality proves Condition 3.5 with $\alpha = 1$.

3.2.3 Verification of Condition 3.6

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 [10]. 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

$$(Q_{J,i}u, v)_J = (u, v)_J \qquad \text{for all } u \in M_J,\ v \in M_{J,i},$$

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

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

$$\mathcal{J}_J = \sum_{i=1}^{N_J} A_{J,i}^{-1} Q_{J,i}. \tag{3.26}$$

The Gauss-Seidel operator on each M_J is defined as

$$\mathcal{G}_J = \big(I - (I - P_{J,N_J})(I - P_{J,N_{J-1}}) \cdots (I - P_{J,1})\big) A_J^{-1}. \tag{3.27}$$

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 norm on $\mathscr{E}_J$.

Lemma 3.12. *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 $\phi_J^{i_1}, \phi_J^{i_2}, \dots \phi_J^{i_{d+1}}$. 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 3.13. *For all v in M_J,*

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

Proof. Since $\mathcal{J}_J$ is an additive operator of the form in (3.26), by a well known lemma on additive operators [10], splitting $v = \sum_{i=1}^{N_J} c_i \phi_J^i$,

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

where we have used Lemma 3.12.

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 [7, 8, 10] the following lemma:

Lemma 3.14. *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*

$$(\tilde{R}_J^{-1}v, v)_J \leq \beta(\mathcal{J}_J^{-1}v, v)_J, \quad \forall v \in M_J.$$

To complete the verification of Condition 3.6 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

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

For $k = J$, the latter inequality follows by combining the estimates of Lemma 3.14 with Lemma 3.13. For the remaining k, the estimate is standard for point Jacobi and Gauss-Seidel operators. Thus Condition 3.6 is verified for all k.

3.3 Numerical Experiments

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

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 [35], 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 3-1. Suppose we need to solve the model problem (2.4) 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 $-\int_{\partial\Omega} g\,\vec{v}\cdot\vec{n}$ on the right hand side of the first equation of (2.7). But the multigrid algorithm is unaffected.

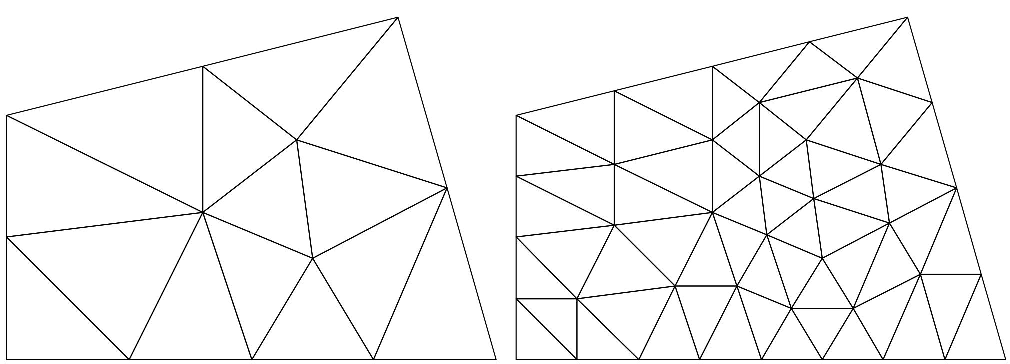

Figure 3-1. Refining mesh: initial mesh on the left, and refined mesh on the right. Corner coordinators in the initial mesh are $(0, 0), (1, 0), (0.8, 0.7)$ and $(0, 0.5)$.

We use two different algorithms to solve the system $Ax = b$, namely the conjugate gradient iteration, and the variable multigrid V-cycle of Algorithm 3.2 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 3-1. 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 $\vec{q}$ as described in § 2.3. Three different kinds of discretization errors are reported in Table 3-2. They show the convergence of finite element errors in accordance with the known theoretical results [2, 18], and as such may be considered to be a validation of our code.

Table 3-1. Performance comparison between unpreconditioned conjugate gradient method and the multigrid method. (Entries marked * indicates unavailable data due to excessive computational time.)

Size	conjugate		multigrid	
	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	9.60
343424	3283	1209.65	34	40.68
1374976	6503	10190.00	34	163.08
5502464	*	*	34	668.63
22014976	*	*	34	2607.65

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 the conjugate gradient method, each time the mesh size is halved. This clearly demonstrates the benefits of the multigrid

Table 3-2. Discretization errors for the hybridized mixed Raviart-Thomas method.

h	$\|Pu - \lambda_h\|_a$	$\|u - u_h\|_{L^2}$	$\|\vec{q} - \vec{q}_h\|_{L^2}$
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

algorithm. Also notice that the number of multigrid iterations seems bounded even as the matrix size gets very large. This is in accordance with the conclusion of Theorem 3.3. In other words, the error reduction factor seems to be independent of mesh size, which is in accordance with the conclusion of Theorem 3.3.

Next, we present an example designed to check if the sufficient condition that Ω is convex (in Theorem 3.3) is necessary. We repeated the experiments with the domain as shown in Figure 3-2. Table 3-3 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 is not convex.

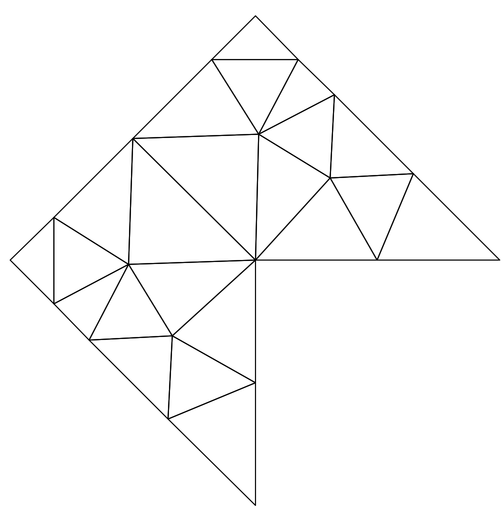

Figure 3-2. The non-convex domain used in experiments.

Table 3-3. Application of the multigrid algorithm to a problem on the non-convex domain of Figure 3-2.

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

Next, we investigate flexibility in regard to the number of smoothings m_k. Theorem 3.3 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 3-4 indicates that the V-cycle algorithm continues to exhibit mesh independent convergence.

Table 3-4. V-cycle with constant number of smoothings.

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

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 3.2 with the nonnested multilevel spaces $M_k = \{\mu \in L^2(\mathscr{E}_k^i) : \mu|_e \in P_0(e), \text{ for all } e \in \mathscr{E}_k^i\}$ and non-inherited forms at every level given by $a_k(u,v) = \int_\Omega \vec{\mathfrak{Q}}^{\mathrm{RT}} u \cdot \vec{\mathfrak{Q}}^{\mathrm{RT}} v$, where now the liftings $\vec{\mathfrak{Q}}^{\mathrm{RT}}(\cdot)$ are defined with respect to $\mathscr{T}_k$. The forms $a_k(\cdot,\cdot)$ and a base inner product as in (3.1) 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 3-3. 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

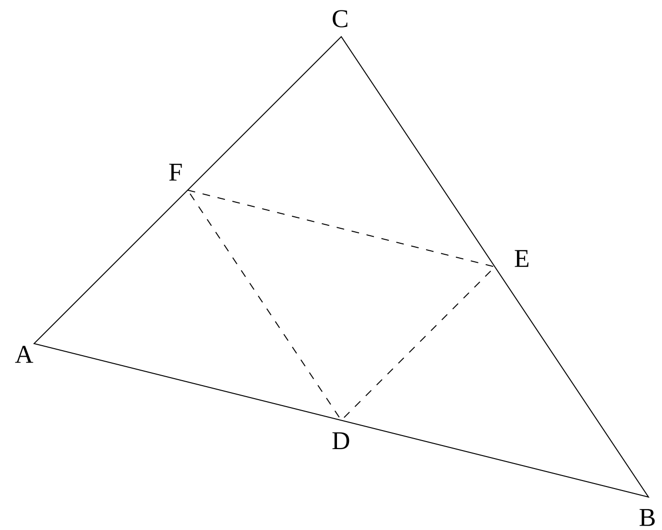

Figure 3-3. A refined triangle with the notations used to describe the failed intergrid transfer operators.

all the finer edges as follows:

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

where $\mathcal{U}^{\mathrm{RT}}(\cdot)$ is the lifting operator defined in (2.8), but now with respect to the mesh $\mathscr{T}_{k-1}$. The second prolongation candidate we shall consider is $I_k^{(2)} : M_{k-1} \to M_k$ is defined by

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

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 3-5. 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

Table 3-5. Failure of certain intergrid operators. (An entry * indicates that the iteration diverged.)

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

finer levels. In contrast, the successful prolongation I_k that we analyzed, does not increase energy, as can be seen from Condition 3.4, which we verified for our I_k.

CHAPTER 4
ESTIMATES FOR THE HDG METHOD

This chapter is devoted to a study of the HDG method introduced in Section 2.4.1. This will form preliminary material for the study of the multigrid methods for HDG methods presented in the next chapter.

HDG methods were discovered in [19]. The first error analysis of such methods was provided in [15]. This analysis was improved using a special projection tailored to the method in [20]. The error estimates we shall present in this chapter are proved using the same projection. These estimates are slightly different from those in [20] (and [15]) in that ours hold under less regularity assumptions. Such low regularity estimates are necessary for the multigrid analysis later, where we do not place convexity or other such assumptions on the domain that guarantee solutions of higher regularity.

Apart from error estimates, we also prove bounds on the spectrum of the operators arising from the HDG method. This yields a condition number estimate for the matrix system resulting from the HDG method, indicating how the performance of classical iterative techniques (like Gauss-Seidel or conjugate gradient methods) deteriorate with system size. They serve as a prelude to the necessity of techniques like multigrid to accelerate solution of HDG methods. Our techniques also easily yield a condition number bound on the HIP method, as we show. While condition number bounds for the HRT methods were known previously, no such bounds were established for the HDG or HIP methods previously.

4.1 Estimates for HDG

The HDG method is previously introduced in § 2.4.1. Its defining equation system is (2.17). In fact, it will be notationally efficient to rewrite this form in terms of the following operators. Define $\mathcal{A} : V_h \mapsto V_h$, $\mathcal{B} : V_h \mapsto W_h$, $\mathcal{C} : V_h \mapsto M_h$, $\mathcal{R} : W_h \mapsto W_h$,

$\mathcal{S} : M_h \mapsto W_h$, and $\mathcal{T} : M_h \mapsto M_h$, by

$$(\mathcal{A}\vec{p}, \vec{v})_{\Omega_h} = (\vec{p}, \vec{v})_{\Omega_h}, \qquad (\mathcal{B}\vec{v}, w)_{\Omega_h} = -(w, \nabla\cdot\vec{v})_{\Omega_h}, \qquad \langle \mathcal{C}\vec{v}, \mu\rangle_{\partial\Omega_h} = \langle \mu, \vec{v}\cdot\vec{n}\rangle_{\partial\Omega_h},$$
$$(\mathcal{R}w, v)_{\Omega_h} = -\langle \tau w, v\rangle_{\partial\Omega_h}, \qquad \langle \mathcal{S}w, \mu\rangle_{\partial\Omega_h} = \langle \tau w, \mu\rangle_{\partial\Omega_h}, \qquad \langle \mathcal{T}\mu, \eta\rangle_{\partial\Omega_h} = -\langle \tau\mu, \eta\rangle_{\partial\Omega_h},$$

for all $\vec{p}, \vec{v} \in V_h$, $w, v \in W_h$, and $\mu, \eta \in M_h$. The HDG method generates operator equations of the form

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^t & \mathcal{C}^t \\ \mathcal{B} & \mathcal{R} & \mathcal{S}^t \\ \mathcal{C} & \mathcal{S} & \mathcal{T} \end{pmatrix} \begin{pmatrix} \vec{q}_h \\ u_h \\ \lambda_h \end{pmatrix} = \begin{pmatrix} \vec{g}_h \\ f_h \\ 0 \end{pmatrix}, \tag{4.1}$$

for some $\vec{g}_h \in V_h$ and $f_h \in W_h$, where the superscript "t" denotes the adjoint with respect to $(\cdot,\cdot)_{\Omega_h}$ or $\langle\cdot,\cdot\rangle_{\partial\Omega_h}$ as appropriate. It is easy to see that (2.15) can be rewritten as the above system with $f_h = \Pi_{W_h} f$, where Π_{W_h} denotes the $L^2(\Omega)$-orthogonal projection into W_h, and $\vec{g}_h$ set to the unique function in V_h satisfying

$$(\vec{g}_h, \vec{v})_{\Omega_h} = -\langle g, \vec{v}\cdot\vec{n}\rangle_{\partial\Omega} \quad \text{for all } \vec{v} \in V_h. \tag{4.2}$$

Note that in the lowest order case $k = 0$, the operator $\mathcal{B}$ is zero, but the system continues to be uniquely solvable.

The result on the above mentioned elimination can be described using additional "local" operators $\vec{\mathcal{Q}}_V^{\mathrm{DG}} : V_h \mapsto V_h$, $\vec{\mathcal{Q}}_W^{\mathrm{DG}} : W_h \mapsto V_h$, $\mathcal{U}_V^{\mathrm{DG}} : V_h \mapsto W_h$, $\mathcal{U}_W^{\mathrm{DG}} : W_h \mapsto W_h$, whose action is defined by solving the following systems

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^t \\ \mathcal{B} & \mathcal{R} \end{pmatrix} \begin{pmatrix} \vec{\mathcal{Q}}_V^{\mathrm{DG}}\vec{g}_h \\ \mathcal{U}_V^{\mathrm{DG}}\vec{g}_h \end{pmatrix} = \begin{pmatrix} \vec{g}_h \\ 0 \end{pmatrix} \qquad \begin{pmatrix} \mathcal{A} & \mathcal{B}^t \\ \mathcal{B} & \mathcal{R} \end{pmatrix} \begin{pmatrix} \vec{\mathcal{Q}}_W^{\mathrm{DG}} f_h \\ \mathcal{U}_W^{\mathrm{DG}} f_h \end{pmatrix} = \begin{pmatrix} 0 \\ f_h \end{pmatrix} \tag{4.3}$$

for all $\vec{g}_h \in V_h$ and $f_h \in W_h$.

One can then readily verify that $\vec{\mathcal{Q}}^{\mathrm{DG}} m = -\vec{\mathcal{Q}}_V^{\mathrm{DG}}(\mathcal{C}^t m) - \vec{\mathcal{Q}}_W^{\mathrm{DG}}(\mathcal{S}^t m)$ and $\mathcal{U}^{\mathrm{DG}} m = -\mathcal{U}_V^{\mathrm{DG}}(\mathcal{C}^t m) - \mathcal{U}_W^{\mathrm{DG}}(\mathcal{S}^t m)$, where $(\vec{\mathcal{Q}}^{\mathrm{DG}} m, \mathcal{U}^{\mathrm{DG}} m)$ is defined in (2.13). For example, (4.3) implies that

$$(\vec{\mathcal{Q}}^{\text{DG}} m, \vec{v})_K - (\mathcal{U}^{\text{DG}} m, \nabla\cdot\vec{v})_K = -\langle m, \vec{v}\cdot\vec{n}\rangle_{\partial K} \qquad \text{for all } \vec{v}\in V_h, \tag{4.4a}$$

$$(w, \nabla\cdot\vec{\mathcal{Q}}^{\text{DG}} m)_K + \langle \tau(\mathcal{U}^{\text{DG}} m - m), w\rangle_{\partial K} = 0 \qquad \text{for all } w\in W_h, \tag{4.4b}$$

which is equivalent to (2.13).

Based on these operators above, we will obtain estimates on the stability, conditioning, and discretization errors of the HDG method. Our technique consists of first obtaining bounds for various local solution operators of the HDG method. The local bounds then imply global bounds, such as bounds for the discretization errors and the length of the spectrum.

4.1.1 Stability of the HDG Local Solvers

In this subsection, we will establish the following result giving bounds on various local solution operators.

Theorem 4.1. *The local solution operators obey the following bounds:*

$$\|\vec{\mathcal{Q}}^{\text{DG}}\mu\|_K \le c^K_{h\tau} C h_K^{-1}\|\mu\|_{h,K}, \qquad \|\mathcal{U}^{\text{DG}}\mu\|_K \le c^K_{h\tau} C\|\mu\|_{h,K}, \tag{4.5}$$

$$\|\vec{\mathcal{Q}}^{\text{DG}}_W f\|_K \le d^K_{h\tau} C h_K\|f\|_K, \qquad \|\mathcal{U}^{\text{DG}}_W f\|_K \le (d^K_{h\tau})^2 C h_K^2\|f\|_K, \tag{4.6}$$

$$\|\vec{\mathcal{Q}}^{\text{DG}}_V \vec{g}\|_K \le C\|\vec{g}\|_K, \qquad \|\mathcal{U}^{\text{DG}}_V \vec{g}\|_K \le d^K_{h\tau} C h_K\|\vec{g}\|_K. \tag{4.7}$$

where $c^K_{h\tau} = 1 + (\tau_K^{\max} h_K)^{1/2}$ *and* $d^K_{h\tau} = 1 + (\tau_K^{\max} h_K)^{-1/2}$ *and* $\tau_K^{\max}$ *denotes the maximum value of* τ_K *on* ∂K*.*

We begin by first defining solution operators $\vec{\mathcal{Q}}^{\text{RT}}_V, \mathcal{U}^{\text{RT}}_V, \vec{\mathcal{Q}}^{\text{RT}}_W, \mathcal{U}^{\text{RT}}_W$ in a similar way as in (4.3), with the obvious modification of the right hand side. It is instructive to compare Theorem 4.1 with a similar result for these HRT operators, as proved in [18, Lemma 3.3]. For instance, one pair of inequalities of [18, Lemma 3.3] is

$$\|\vec{\mathcal{Q}}^{\text{RT}}\mu\|_K \le C h_K^{-1}\|\mu\|_{h,K}, \qquad \|\mathcal{U}^{\text{RT}}\mu\|_K \le C\|\mu\|_{h,K}, \tag{4.8}$$

which is comparable to (4.5), when τ is of unit size. More interestingly, cf. (4.6) with

$$\|\vec{\mathcal{Q}}_W^{\text{RT}} f\|_K \le C h_K \|f\|_K, \qquad\qquad \|\mathcal{U}_W^{\text{RT}} f\|_K \le C h_K^2 \|f\|_K,$$

which is another pair of inequalities of [18, Lemma 3.3]. Observe that if τ is of unit size, these local HRT operators are more stable than the corresponding HDG ones. Indeed, while $\mathcal{U}_W^{\text{RT}} f$ damps perturbations in f by $O(h^2)$, the corresponding HDG operator, namely $\mathcal{U}_W^{\text{DG}} f$, damps it by only $O(h)$ because $(d_{h\tau}^K)^2 = O(1/h)$.

We will now develop a series of intermediate results to prove Theorem 4.1 in the remainder of this subsection.

Lemma 4.2. *For all λ in M_h,*

$$\|\mathcal{U}^{\text{RT}} \lambda - \lambda\|_{\partial K} \le C h_K^{1/2} \|\vec{\mathcal{Q}}^{\text{RT}} \lambda\|_K.$$

Proof. Integrating (2.8) by parts on a mesh element K with $m = \lambda$, we have

$$(\vec{\mathcal{Q}}^{\text{RT}} \lambda + \vec{\nabla}\, \mathcal{U}^{\text{RT}} \lambda, \vec{v})_K = \langle \mathcal{U}^{\text{RT}} \lambda - \lambda, \vec{v} \cdot \vec{n} \rangle_{\partial K}. \tag{4.9}$$

There is an $\vec{v}$ in $R_d(K)$ such that $\vec{v} \cdot \vec{n} = \mathcal{U}^{\text{RT}} \lambda - \lambda$ on ∂K and $(\vec{v}, \vec{p}_{d-1})_K = 0$ for all $\vec{p}_{d-1}$ in $P_{d-1}(K)^n$ (this is obvious from the well-known degrees of freedom of the space $R_d(K)$). Additionally, by a scaling argument it is immediate that

$$\|\vec{v}\|_K \le C h_K^{1/2} \|\vec{v} \cdot \vec{n}\|_{\partial K}. \tag{4.10}$$

With this $\vec{v}$ in (4.9), we obtain

$$\|\mathcal{U}^{\text{RT}} \lambda - \lambda\|_{\partial K}^2 = (\vec{\mathcal{Q}}^{\text{RT}} \lambda, \vec{v})_K,$$

from which the lemma follows by Cauchy-Schwarz inequality and (4.10).

Lemma 4.3. *If F is any face of the simplex K,*

$$C \|w\|_K \le h_K \|\mathcal{B}^t w\|_K + h_K^{1/2} \|w\|_F, \qquad \forall\, w \in W_h.$$

Proof. On the unit simplex $\hat{K}$, we have

$$\hat{C}\|\hat{w}\|_{\hat{K}} \leq \sup_{\vec{v}\in P_d(\hat{K})} \frac{|(\hat{w}, \nabla\cdot\vec{v})_{\hat{K}}|}{\|\vec{v}\,\|_{\hat{K}}} + \|\hat{w}\|_{\hat{F}}, \qquad \forall\, \hat{w} \in P_d(K), \tag{4.11}$$

for any face $\hat{F}$ of $\hat{K}$. This follows by equivalence of norms. That the right hand side indeed defines a norm can be seen as follows: divergence is a surjective map from $P_d(K)^n$ to $P_{d-1}(K)$. Hence if the supremum is zero, then $\hat{w}$ is orthogonal to $P_{d-1}(K)$, in which case $\hat{w}$ is zero once it vanishes on any face $\hat{F}$ (see [20, Lemma 2.1]). The lemma follows by mapping (4.11) to any simplex K and using standard scaling arguments.

Lemma 4.4. *Let μ be any function in M_h. The following statements hold:*

(i) *If $\mu|_{\partial K} = v|_{\partial K}$ for some $v \in P_0(K)$, then*

$$\mathfrak{U}^{\mathrm{DG}}\mu|_{\partial K} = \mu|_{\partial K} \qquad \textit{and} \qquad \vec{\mathfrak{Q}}^{\mathrm{DG}}\mu = 0.$$

(ii) *If $d > 0$, and $\mu|_{\partial K} = v|_{\partial K}$ for some $v \in P_1(K)$, then*

$$\mathfrak{U}^{\mathrm{DG}}\mu = v \qquad \textit{and} \qquad \vec{\mathfrak{Q}}^{\mathrm{DG}}\mu = -\vec{\nabla}\, v.$$

(iii) *We have the following bounds (where $\|\mu\|_{\tau,\partial K} = \langle\tau\mu,\mu\rangle_{\partial K}^{1/2}$):*

$$\|\mathfrak{U}^{\mathrm{DG}}\mu - \mu\|_{\tau,\partial K} \leq C\sqrt{\tau_K^{\max} h_K}\,\|\vec{\mathfrak{Q}}^{\mathrm{RT}}\mu\|_K, \tag{4.12}$$

$$\|\vec{\mathfrak{Q}}^{\mathrm{DG}}\mu - \vec{\mathfrak{Q}}^{\mathrm{RT}}\mu\|_K \leq C\sqrt{\tau_K^{\max} h_K}\,\|\vec{\mathfrak{Q}}^{\mathrm{RT}}\mu\|_K, \tag{4.13}$$

$$\|\mathfrak{U}^{\mathrm{DG}}\mu - \mathfrak{U}^{\mathrm{RT}}\mu\|_K \leq C h_K\,(1 + \sqrt{\tau_K^{\max} h_K})\,\|\vec{\mathfrak{Q}}^{\mathrm{RT}}\mu\|_K. \tag{4.14}$$

(iv) *If J_K is the $L^2(K)-$orthogonal projection into $\{\vec{v} \in P_d(K)^n : \nabla\cdot\vec{v} = 0\}$ then*

$$\vec{\mathfrak{Q}}^{\mathrm{RT}}\mu = J_K\vec{\mathfrak{Q}}^{\mathrm{DG}}\mu. \tag{4.15}$$

In particular,

$$\|\vec{\mathfrak{Q}}^{\mathrm{RT}}\mu\|_K \leq \|\vec{\mathfrak{Q}}^{\mathrm{DG}}\mu\|_K. \tag{4.16}$$

Proof. This proof proceeds by comparing the RT and HDG equations for the local solutions. Subtracting (2.8) from (2.13) we have

$$((\vec{\mathfrak{Q}}^{\text{DG}}\mu - \vec{\mathfrak{Q}}^{\text{RT}}\mu), \vec{v})_K - (\mathcal{U}^{\text{DG}}\mu - \mathcal{U}^{\text{RT}}\mu, \nabla\cdot\vec{v})_K = 0 \tag{4.17a}$$

$$(\nabla\cdot(\vec{\mathfrak{Q}}^{\text{DG}}\mu - \vec{\mathfrak{Q}}^{\text{RT}}\mu), w)_K + \langle\tau(\mathcal{U}^{\text{DG}}\mu - \mathcal{U}^{\text{RT}}\mu), w\rangle_{\partial K} = \langle\tau(\mu - \mathcal{U}^{\text{RT}}\mu), w\rangle_{\partial K} \tag{4.17b}$$

for all $\vec{v} \in P_d(K)^n$ and for all w in $P_d(K)$. Note that since $\nabla\cdot\vec{\mathfrak{Q}}^{\text{RT}}\mu = 0$, the lifting $\vec{\mathfrak{Q}}^{\text{RT}}\mu$ is in fact in $P_d(K)^n$. Hence $\{\vec{\mathfrak{Q}}^{\text{DG}}\mu - \vec{\mathfrak{Q}}^{\text{RT}}\mu, \mathcal{U}^{\text{DG}}\mu - \mathcal{U}^{\text{RT}}\mu\}$ forms the unique solution of (4.17).

First, let us prove the first assertion (i) of the lemma. Indeed, if μ takes a constant value on ∂K, then it is well known that $\mathcal{U}^{\text{RT}}\mu$ equals the same constant [17] and $\vec{\mathfrak{Q}}^{\text{RT}}\mu = 0$, so the right hand side of (4.17) vanishes. Hence $\mathcal{U}^{\text{RT}}\mu - \mathcal{U}^{\text{DG}}\mu$ and $\vec{\mathfrak{Q}}^{\text{RT}}\mu - \vec{\mathfrak{Q}}^{\text{DG}}\mu$ also vanish, thus proving (i).

The statement (ii) is proved by the same technique as (i). The only difference is that the analogous result for the RT case is less well known, so let us first show it, namely $\mathcal{U}^{\text{RT}}\mu|_{\partial K} = \mu|_{\partial K}$ when $d > 0$ and $\mu|_{\partial K}$ equals the trace of some $v \in P_1(K)$. In light of (3.7), equation (2.8) becomes

$$-(\vec{\nabla}\, v, \vec{v}) - (\mathcal{U}^{\text{RT}}\mu, \nabla\cdot\vec{v}) = -\langle v, \vec{v}\cdot\vec{n}\rangle_{\partial K} = -(\vec{\nabla}\, v, \vec{v}) - (v, \nabla\cdot\vec{v}).$$

This implies that

$$(\mathcal{U}^{\text{RT}}\mu - v, \nabla\cdot\vec{v}) = 0 \qquad \forall\vec{v} \in R_d(K),$$

so that $\mathcal{U}^{\text{RT}}\mu = v$. Thus, just as in the proof of item (i), the solution of (4.17) vanishes in this case also, and we have proven item (ii).

Next, let us prove the estimates. Setting $\vec{v} = \vec{\mathfrak{Q}}^{\text{DG}}\mu - \vec{\mathfrak{Q}}^{\text{RT}}\mu$ and $w = \mathcal{U}^{\text{DG}}\mu - \mathcal{U}^{\text{RT}}\mu$ in (4.17), we have

$$\|\vec{\mathfrak{Q}}^{\text{DG}}\mu - \vec{\mathfrak{Q}}^{\text{RT}}\mu\|_K^2 + \|\mathcal{U}^{\text{DG}}\mu - \mathcal{U}^{\text{RT}}\mu\|_{\tau,\partial K}^2 = \langle\tau(\mu - \mathcal{U}^{\text{RT}}\mu), \mathcal{U}^{\text{DG}}\mu - \mathcal{U}^{\text{RT}}\mu\rangle_{\partial K}. \tag{4.18}$$

The identity (4.18) implies

$$\|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{\tau,\partial K} \leq \|\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{\tau,\partial K},$$

from which (4.12) follows by triangle inequality and Lemma 4.2.

Applying Cauchy-Schwarz inequality to the right hand side of (4.18),

$$\|\vec{\mathcal{Q}}^{\mathrm{DG}}\mu - \vec{\mathcal{Q}}^{\mathrm{RT}}\mu\|_K^2 + \|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{\tau,\partial K}^2 \leq C h_K^{1/2} \|\vec{\mathcal{Q}}^{\mathrm{RT}}\mu\|_K \tau_K^{\max 1/2} \|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{\tau,\partial K}$$

by Lemma 4.2. Clearly (4.13) immediately follows.

It remains only to prove (4.14). Let $F_{\max}$ denote a face of K where $\tau = \tau_K^{\max}$. Then

$$\tau_K^{\max} \|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{F_{\max}}^2 = \|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{\tau,F_{\max}}^2 \leq \|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{\tau,\partial K}^2 \leq C\tau_K^{\max} h_K \|\vec{\mathcal{Q}}^{\mathrm{RT}}\mu\|_K^2,$$

so canceling off the common factor $\tau_K^{\max}$, we have

$$\|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{F_{\max}} \leq C h_K^{1/2} \|\vec{\mathcal{Q}}^{\mathrm{RT}}\mu\|_K.$$

Hence using Lemma 4.3, we obtain

$$\begin{aligned}
\|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_K &\leq C \left(h_K \|\mathcal{B}^t(\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu)\|_K + h_K^{1/2} \|\mathcal{U}^{\mathrm{DG}}\mu - \mathcal{U}^{\mathrm{RT}}\mu\|_{F_{\max}} \right) \\
&\leq C \left(h_K \|\vec{\mathcal{Q}}^{\mathrm{DG}}\mu - \vec{\mathcal{Q}}^{\mathrm{RT}}\mu\|_K + h_K \|\vec{\mathcal{Q}}^{\mathrm{RT}}\mu\|_K \right),
\end{aligned}$$

from which (4.14) follows. (This applies even if $d = 0$, in which case the term involving $\mathcal{B}^t$ is absent.) Thus we have proved item (iii).

For the final item (iv),

$$\begin{aligned}
(J_K \vec{\mathcal{Q}}^{\mathrm{DG}}\mu, J_K \vec{v})_K &= -\langle \mu, J_K \vec{v} \cdot \vec{n} \rangle_{\partial K} && \text{by (2.13)}, \\
&= (\vec{\mathcal{Q}}^{\mathrm{RT}}\mu, J_K \vec{v})_K, && \text{by (2.8)},
\end{aligned}$$

which proves the equality (4.15), as $\vec{\mathcal{Q}}^{\mathrm{RT}}\mu$ is in the range of J_K by (2.8). The estimate (4.16) is then obvious as orthogonal projectors have unit norm.

Proof of Theorem 4.1. First, we prove the bounds on $\vec{\mathfrak{Q}}^{\text{DG}}\mu, \mathfrak{U}^{\text{DG}}\mu$:

$$\begin{aligned}
\|\vec{\mathfrak{Q}}^{\text{DG}}\mu\|_K &\le \|\vec{\mathfrak{Q}}^{\text{RT}}\mu\|_K + \|\vec{\mathfrak{Q}}^{\text{DG}}\mu - \vec{\mathfrak{Q}}^{\text{RT}}\mu\|_K \\
&\le \|\vec{\mathfrak{Q}}^{\text{RT}}\mu\|_K + C(\tau_K^{\max} h_K)^{1/2}\|\vec{\mathfrak{Q}}^{\text{RT}}\mu\|_K && \text{by (4.13) of Lemma 4.4,} \\
&\le c_{h\tau}^K C\|\mu\|_{h,K} && \text{by (4.8).}
\end{aligned}$$

The bound for $\mathfrak{U}^{\text{DG}}\mu$ is proved similarly using (4.13) in place of (4.13).

Next, consider $\vec{\mathfrak{Q}}_W^{\text{DG}} f, \mathfrak{U}_W^{\text{DG}} f$. From their definitions, it is easy to see that

$$(\vec{\mathfrak{Q}}_W^{\text{DG}} f, \vec{\mathfrak{Q}}_W^{\text{DG}} f)_K + \langle \tau \mathfrak{U}_W^{\text{DG}} f, \mathfrak{U}_W^{\text{DG}} f\rangle_{\partial K} = (\mathfrak{U}_W^{\text{DG}} f, f)_K. \tag{4.19}$$

Let $F_{\max}$ denote a face of K where $\tau = \tau_K^{\max}$. Then,

$$\begin{aligned}
\|\mathfrak{U}_W^{\text{DG}} f\|_K &\le C(h_K\|\mathcal{B}^t\mathfrak{U}_W^{\text{DG}} f\|_K + h_K^{1/2}\|\mathfrak{U}_W^{\text{DG}} f\|_{F_{\max}}) && \text{by Lemma 4.3,} \\
&\le C(h_K\|\vec{\mathfrak{Q}}_W^{\text{DG}} f\|_K + h_K^{1/2}\|\mathfrak{U}_W^{\text{DG}} f\|_{F_{\max}}) && \text{as } \mathcal{A}\vec{\mathfrak{Q}}_W^{\text{DG}} f + \mathcal{B}^t\mathfrak{U}_W^{\text{DG}} f = 0, \\
&\le Ch_K(\|\vec{\mathfrak{Q}}_W^{\text{DG}} f\|_K + (\tau_K^{\max} h_K)^{-1/2}\|\mathfrak{U}_W^{\text{DG}} f\|_{\tau,\partial K}) \\
&\le Ch_K((\mathfrak{U}_W^{\text{DG}} f, f)_K^{1/2} + (\tau_K^{\max} h_K)^{-1/2}(\mathfrak{U}_W^{\text{DG}} f, f)_K^{1/2}) && \text{by (4.19),} \\
&\le Ch_K(1 + (\tau_K^{\max} h_K)^{-1/2})\|\mathfrak{U}_W^{\text{DG}} f\|_K^{1/2}\|f\|_K^{1/2},
\end{aligned}$$

from which the required bound on $\mathfrak{U}_W^{\text{DG}} f$ follows. Using this in (4.19), we immediately get the stated bound for $\vec{\mathfrak{Q}}_W^{\text{DG}} f$ as well.

Finally, to prove (4.7), we start from the following easy consequence of the definitions of $\vec{\mathfrak{Q}}_V^{\text{DG}}\vec{g}, \mathfrak{U}_V^{\text{DG}}\vec{g}$:

$$\|\vec{\mathfrak{Q}}_V^{\text{DG}}\vec{g}\|_K^2 + \|\mathfrak{U}_V^{\text{DG}}\vec{g}\|_{\tau,\partial K}^2 = (\vec{g}, \vec{\mathfrak{Q}}_V^{\text{DG}}\vec{g})_K. \tag{4.20}$$

Since it is immediate from the above that $\|\vec{\mathfrak{Q}}_V^{\text{DG}}\vec{g}\|_K \le C\|\vec{g}\|_K$, it only remains to prove the bound for $\mathfrak{U}_V^{\text{DG}}\vec{g}$. By Lemma 4.3,

$$\begin{aligned}
\|\mathfrak{U}_V^{\text{DG}}\vec{g}\|_K &\le Ch_K\|\mathcal{B}^t\mathfrak{U}_V^{\text{DG}}\vec{g}\|_K + Ch_K^{1/2}\|\mathfrak{U}_V^{\text{DG}}\vec{g}\|_{F_{\max}} \\
&\le Ch_K(\|\mathcal{A}\vec{\mathfrak{Q}}_V^{\text{DG}}\vec{g}\|_K + \|\vec{g}\|_K) + Ch_K(\tau_K^{\max} h_K)^{-1/2}\|\vec{\mathfrak{Q}}_V^{\text{DG}}\vec{g}\|_{\tau,\partial K},
\end{aligned}$$

and the final inequality of the theorem follows by using (4.20) in the above.

4.1.2 Error Estimates for the HDG Method

Error estimates for the HDG method under consideration have been proved in [20]. Here, as an application of the estimates we proved in § 4.1.1, we prove two new error estimates not in [20].

The proof is quick and easy once we use the special projection of [20]. The projection, denoted by $\mathit{\Pi}_h(\vec{q}, u)$, is into the product space $V_h \times W_h$, and its domain is a subspace of $H(\mathrm{div}, \Omega) \times L^2(\Omega)$ consisting of sufficiently regular functions, e.g.,

$$H(\mathrm{div}, \Omega) \cap H^s(\Omega)^n \times H^s(\Omega) \qquad \text{for } s > 1/2.$$

When its components need to be identified, we also write $\mathit{\Pi}_h(\vec{q}, u)$ as $(\mathit{\Pi}_h^V \vec{q}, \mathit{\Pi}_h^W u)$ where $\mathit{\Pi}_h^V \vec{q}$ and $\mathit{\Pi}_h^W u$ are the components of the projection in V_h and W_h, respectively. (Despite this notation, note that $\mathit{\Pi}_h^V \vec{q}$ depends not just on $\vec{q}$, but rather on both $\vec{q}$ and u. The same applies for $\mathit{\Pi}_h^W u$.) We omit the definition and other details of the projection, but we need the following approximation property proved in [20]:

$$\|\vec{q} - \mathit{\Pi}_h^V \vec{q}\|_K + \tau_K^{\max} \|u - \mathit{\Pi}_h^W u\|_K \leq C h^s \left(|\vec{q}|_{H^s(K)} + \tau_K^{\max} |u|_{H^s(K)} \right) \tag{4.21}$$

for all $1/2 < s \leq k + 1$.

When the approximation property (4.21) is combined with the following theorem, we get optimal estimates for all variables of the HDG method. The additional notations used in the theorem are as follows: P_h^M denotes the orthogonal projection into M_h defined by

$$\langle P_h^M u, \mu \rangle_{\partial\Omega_h} = \langle u, \mu \rangle_{\partial\Omega_h}, \quad \text{for all } \mu \in M_h.$$

Let $\|\mu\|_a = a_h(\mu, \mu)^{1/2}$ and $\|\vec{v}\,\| = (\vec{v}, \vec{v})^{1/2}$.

Theorem 4.5. *Let the exact solution satisfying* (2.12) *be* $(\vec{q}, u)$*, and the discrete solution satisfying* (2.17) *be* $(\vec{q}_h, u_h, \lambda_h)$*. Then, the following error estimates hold:*

$$\|\vec{q} - \vec{q}_h\| \leq 2\|\vec{q} - \Pi_h^V \vec{q}\|, \tag{4.22}$$

$$\|P_h^M u - \lambda_h\|_a \leq \ \|\vec{q} - \Pi_h^V \vec{q}\|, \tag{4.23}$$

$$\|u - u_h\|_{\Omega_h} \leq C\|u - \Pi_h^W u\| + b_\tau C\|\vec{q} - \Pi_h^V \vec{q}\|, \tag{4.24}$$

where $b_\tau = \max\{1 + h_K \tau_K^{\max} + h_K/\tau_K^{\max} : K \in \mathscr{T}_h\}$.

The main tool for proving these error estimates is a system of equations that shows that $\Pi_h^V \vec{q}, \Pi_h^W u$, and $P_h^M u$ satisfies the same equations as $\vec{q}_h, u_h$, and λ_h up to a perturbation in the right hand side. This is seen in the next lemma, which is proved in [20]:

Lemma 4.6. *Let* $\vec{\varepsilon}_h^{\,q} = \Pi_h^V \vec{q} - \vec{q}_h$, $\varepsilon_h^u = \Pi_h^W u - u_h$, *and* $\varepsilon_h^\lambda = P_h^M u - \lambda_h$*. Then*

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^t & \mathcal{C}^t \\ \mathcal{B} & \mathcal{R} & \mathcal{S}^t \\ \mathcal{C} & \mathcal{S} & \mathcal{T} \end{pmatrix} \begin{pmatrix} \vec{\varepsilon}_h^{\,q} \\ \varepsilon_h^u \\ \varepsilon_h^\lambda \end{pmatrix} = \begin{pmatrix} \vec{e}_h \\ 0 \\ 0 \end{pmatrix}, \tag{4.25}$$

where $\vec{e}_h$ *is the unique function in* V_h *satisfying* $(\vec{e}_h, \vec{v})_{\Omega_h} = ((\Pi_h^V \vec{q} - \vec{q}\,), \vec{v})_{\Omega_h}$.

Proof. See [20, Lemma 3.1].

Proof of Theorem 4.5. The first estimate is proved in [20] (and the proof is easy), so we only prove the remaining two.

To prove (4.23), we apply Theorem 2.10 to (4.25) of Lemma 4.6. Then, we find that ε_h^λ satisfies

$$a_h(\varepsilon_h^\lambda, \mu) = (-\vec{e}_h, \vec{\mathcal{Q}}^{\mathrm{DG}} \mu) = ((\Pi_h^V \vec{q} - \vec{q}\,), \vec{\mathcal{Q}}^{\mathrm{DG}} \mu)$$

for all μ in M_h. Hence (4.23) follows by choosing $\mu = \varepsilon_h^\lambda$ and applying the Cauchy-Schwarz inequality.

To prove (4.24), we apply the local recovery equation (2.16) of Theorem 2.10 to (4.25), which gives $\varepsilon_h^u = \mathcal{U}^{\mathrm{DG}}\varepsilon_h^\lambda + \mathcal{U}_V^{\mathrm{DG}}\vec{e}_h$. Therefore,

$$\|\varepsilon_h^u\|_K \le \|\mathcal{U}^{\mathrm{DG}}\varepsilon_h^\lambda\|_K + \|\mathcal{U}_V^{\mathrm{DG}}\vec{e}_h\|_K \le c_{h\tau}^K C\|\varepsilon_h^\lambda\|_{h,K} + C d_{h\tau}^K h_K \|\vec{e}_h\|_K,$$

where we have used Theorem 4.1. Since $d_{h\tau}^K h_K \le C h_K^{1/2}(\tau_K^{\max})^{-1/2}$,

$$\|u-u_h\|_K^2 \le C(c_{h\tau}^K)^2\|\varepsilon_h^\lambda\|_{h,K}^2 + C(\tau_K^{\max})^{-1}h_K\|\mathit{\Pi}_h^V\vec{q}-\vec{q}\,\|_K^2 + C\|u-\mathit{\Pi}_h^W u\|_K^2.$$

Summing over all mesh elements and using Lemma 4.9, we obtain

$$\|u-u_h\|_{\Omega_h}^2 \le b_\tau^{(1)} C\|\varepsilon_h^\lambda\|_a^2 + C b_\tau^{(2)}\|\mathit{\Pi}_h^V\vec{q}-\vec{q}\|_{\Omega_h}^2 + C\|u-\mathit{\Pi}_h^W u\|_{\Omega_h}^2.$$

where $b_\tau^{(1)} = \max\{1+\tau_K^{\max}h_K : K \in \mathscr{T}_h\}$ and $b_\tau^{(2)} = \max\{h_K/\tau_K^{\max} : K \in \mathscr{T}_h\}$. Thus we can finish the proof of (4.24) using the previous estimate (4.23) for ε_h^λ.

Remark 4.7. Stronger error estimates for u_h and λ_h are established in [20] under additional regularity assumptions. The only regularity requirement for the estimates (4.24) and (4.23) to hold is that $(\vec{q}, u)$ is in the domain of $\mathit{\Pi}_h$, whereas the analysis in [20] assumes in addition the full regularity condition suitable for an Aubin-Nitsche type argument.

4.1.3 Conditioning of the HDG Method

We now obtain bounds on the spectrum of the operator generated by $a_h(\cdot,\cdot)$. It turns out that the condition number for the HDG method has a similar estimate as for the HRT method (see Theorem 3.1). The main result of this subsection is the following.

Theorem 4.8. *Suppose $\mathscr{T}_h$ is quasiuniform and $h = \max(h_K : K \in \mathscr{T}_h)$. There are positive constants C_1 and C_2 independent of h such that*

$$C_1\|\mu\|_h^2 \le a_h(\mu,\mu) \le \gamma_{h\tau}^{(2)} C_2 h^{-2}\|\mu\|_h^2, \quad \textit{for all } \mu \in M_h \tag{4.26}$$

where $\gamma_{h\tau}^{(2)} = \max\{1+(\tau_K^{\max}h_K)^2 : K \in \mathscr{T}_h\}$.

The implication of this theorem for a condition number bound is as follows. Consider the stiffness matrix of $a_h(\cdot,\cdot)$, obtained through any standard local (face by face) finite element basis for M_h. Let κ be the spectral condition number of this stiffness matrix. Then standard arguments using the two-sided estimate of Theorem 4.8 imply

$$\kappa \le \gamma_{h\tau}^{(2)} C h^{-2}.$$

In particular, note that for all choices of τ satisfying

$$\tau \le C/h,$$

the condition number grows like $O(h^{-2})$. (For the so-called "super-penalized" cases where τ is chosen to be $O(1/h^{\alpha})$ with $\alpha > 1$, it grows even faster.) Again, the growth of the condition number implies a deterioration in the performance of many iterative techniques as h decreases, and hence motivates our development of efficient multigrid algorithms that converge at an h-independent rate.

The proof of Theorem 4.8 relies on the two lemmas below. To state them, we need an additional norm previously defined in Notation 2.9. Note that since $\|\cdot\|_h$ is an L^2-like norm, $|\!|\!|\cdot|\!|\!|_h$ is an H^1-like norm, and since functions in M_h can be thought of as having zero boundary conditions on $\partial\Omega$, it is not surprising that the following Poincaré -type inequality holds:

Lemma 4.9. *There is a constant C_0 such that on all quasiuniform meshes*

$$\|\mu\|_h \le C_0 |\!|\!|\mu|\!|\!|_h \qquad \text{for all } \mu \in M_h. \tag{4.27}$$

Proof. See [24, Proof of Theorem 2.3].

Lemma 4.10. *For all μ in M_h and all mesh elements K,*

$$C |\!|\!|\mu|\!|\!|_{h,K} \le \|\vec{\mathcal{Q}}^{\mathrm{DG}}\mu\|_K \tag{4.28}$$

Proof. If we use the inequality $\|\!|\lambda|\!\|_{h,K} \leq C\|\vec{\mathfrak{Q}}^{\text{RT}}\lambda\|_K$ established in [24], the proof of lemma can be completed instantly by

$$\|\!|\lambda|\!\|_{h,K} \leq C\|\vec{\mathfrak{Q}}^{\text{RT}}\lambda\|_K \leq C\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_K$$

where we have used (4.16) of Lemma 4.4.

Remark 4.11 (Another proof). To give a better idea of how the $\|\!|\cdot|\!\|_{h,K}$-norm enters the arena, we present a second argument below, which is more direct.

Let T_K be the affine isomorphism mapping the reference unit simplex $\hat{K}$ one-one onto K. It has the form $T_K(\hat{x}) = M_K\hat{x} + b$ for some $n \times n$ matrix M_K. We will also need the Piola map Φ_K mapping functions on K to $\hat{K}$, defined by $\Phi_K(\vec{v}) = (\det M_K)^{-1}M_K^{-1}\vec{v} \circ T_K$. We start by letting $\hat{\lambda} = \lambda \circ T_K$ and recalling that there is a function $\vec{v}_{\hat{\lambda}}$ in $P_k(K)^n$ such that

$$\begin{aligned}
&\nabla \cdot \vec{v}_{\hat{\lambda}} = 0, && \text{in } \hat{K},\\
&\vec{v}_{\hat{\lambda}} \cdot \vec{n} = \hat{\lambda} - m_{\hat{K}}(\hat{\lambda}), && \text{on } \partial\hat{K}, \text{ and}\\
&\|\vec{v}_{\hat{\lambda}}\|_{\hat{K}} \leq C\|\hat{\lambda} - m_{\hat{K}}(\hat{\lambda})\|_{\hat{K}}.
\end{aligned}$$

Such an $\vec{v}_{\hat{\lambda}}$ can be obtained, e.g., by the polynomial extension in [13] applied to $\hat{\lambda} - m_{\hat{K}}(\hat{\lambda})$, or even by more elementary observations. Next let $\vec{v}_\lambda = \Phi_K^{-1}(\vec{v}_{\hat{\lambda}})$. By the well known properties of the Piola map [13], we know that

$$\nabla \cdot \vec{v}_\lambda = 0 \quad \text{and} \quad \langle \lambda, \vec{v}_\lambda \cdot \vec{n}\rangle_{\partial K} = \langle \hat{\lambda}, \vec{v}_{\hat{\lambda}} \cdot \vec{n}\rangle_{\partial \hat{K}}.$$

Setting $\vec{v}$ equal to $\vec{v}_\lambda$ in (4.4a), we get

$$\begin{aligned}
(\vec{\mathfrak{Q}}^{\text{DG}}\lambda, \vec{v}_\lambda)_K &= -\langle \lambda, \vec{v}_\lambda \cdot \vec{n}\rangle_{\partial K} = -\langle \hat{\lambda}, \vec{v}_{\hat{\lambda}} \cdot \vec{n}\rangle_{\partial\hat{K}}\\
&= -(\hat{\lambda} - m_{\hat{K}}(\hat{\lambda}), \vec{v}_{\hat{\lambda}} \cdot \vec{n})_{\partial\hat{K}}\\
&= -\|\hat{\lambda} - m_{\hat{K}}(\hat{\lambda})\|^2_{\partial\hat{K}}.
\end{aligned}$$

This implies

$$\|\hat{\lambda} - m_{\hat{K}}(\hat{\lambda})\|^2_{\partial\hat{K}} \leq \|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|_K \, \|\vec{v}_\lambda\|_K \leq C\|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|_K \, \|\vec{v}_{\hat{\lambda}}\|_{\hat{K}} \leq \|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|_K \, \|\hat{\lambda} - m_{\hat{K}}(\hat{\lambda})\|_{\partial\hat{K}},$$

so

$$\|\hat{\lambda} - m_{\hat{K}}(\hat{\lambda})\|_{\partial\hat{K}} \leq C\|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|_K.$$

Using the fact that $m_K(\lambda)$ is the best approximating constant on ∂K to λ, and using a scaling argument,

$$\|\lambda - m_K(\lambda)\|_{\partial K} \leq \|\lambda - m_{\hat{K}}(\hat{\lambda})\|_{\partial K} \leq C h_K^{1/2}\|\hat{\lambda} - m_{\hat{K}}(\hat{\lambda})\|_{\partial\hat{K}}.$$

Therefore,

$$|\!|\!|\lambda|\!|\!|_{h,K} = C h_K^{-1/2}\|\lambda - m_K(\lambda)\|_{\partial K} \ \leq C\|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|_K$$

and the lemma is proved.

Proof of Theorem 4.8. For the upper bound, we use (4.12) and (4.16) of Lemma 4.4 to conclude that

$$\|\mathcal{U}^{\rm DG}\lambda - \lambda\|^2_{\tau,\partial K} \leq C\tau_K^{\max} h_K\|\vec{\mathfrak{Q}}^{\rm RT}\lambda\|^2_K \leq C\tau_K^{\max} h_K\|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|^2_K.$$

Hence, summing over all elements, and denoting $\|\cdot\|_\tau = \langle\tau\cdot,\cdot\rangle^{1/2}_{\partial\Omega_h}$,

$$\begin{aligned} a_h(\lambda,\lambda) = \|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|^2 + \|\mathcal{U}^{\rm DG}\lambda - \lambda\|^2_\tau &\leq C\sum_{K\in\mathscr{T}_h}(1+\tau_K^{\max}h_K)\|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|^2_K \\ &\leq C\sum_{K\in\mathscr{T}_h}(1+\tau_K^{\max}h_K)(c^K_{h\tau})^2 h_K^{-2}\|\lambda\|^2_h, \end{aligned}$$

where we have used Theorem 4.1. Thus, the upper bound follows.

For the lower bound, we combine the estimates of Lemmas 4.9 and 4.10 to obtain

$$\|\lambda\|^2_h \leq C_0|\!|\!|\lambda|\!|\!|^2_h \leq C_0 C\|\vec{\mathfrak{Q}}^{\rm DG}\lambda\|^2_{\Omega_h} \leq CC_0\, a_h(\lambda,\lambda),$$

so the proof is complete.

Remark 4.12. We provide another proof for estimating the upper bound as follows:

Setting $m = \lambda, \vec{v} = \vec{\mathfrak{Q}}^{\text{DG}}\lambda, w = \mathfrak{U}^{\text{DG}}\lambda$ in (2.13), we have

$$a_h^{\text{DG}}(\lambda,\lambda)|_K = \|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_K + \|\mathfrak{U}^{\text{DG}}\lambda - \lambda\|_{\tau,\partial K} = -(\lambda, \vec{\mathfrak{Q}}^{\text{DG}}\lambda \cdot \vec{n})_{\partial K} - (\lambda, \tau(\mathfrak{U}^{\text{DG}}\lambda - \lambda))_{\partial K}.$$

Bound the terms on the right side, and we have

$$|(\lambda, \vec{\mathfrak{Q}}^{\text{DG}}\lambda \cdot \vec{n})_{\partial K}| \leq \|\lambda\|_{\partial K}\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_{\partial K} \leq h^{-1/2}(\|\lambda\|_{\partial K}h^{1/2})(Ch^{-1/2}\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_K)$$

where the last inequality comes from a scaling argument, and

$$|(\lambda, \tau(\mathfrak{U}^{\text{DG}}\lambda - \lambda))_{\partial K}| \leq (\tau_K^{\max})^{1/2}h^{-1/2}(\|\lambda\|_{\partial K}h^{1/2})\|\tau^{1/2}(\mathfrak{U}^{\text{DG}}\lambda - \lambda)\|_{\partial K}.$$

Hence

$$\begin{aligned}
&\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_K + \|\mathfrak{U}^{\text{DG}}\lambda - \lambda\|_{\tau,\partial K} \\
&\leq Ch^{-1}(\|\lambda\|_{\partial K}h^{1/2})(\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_K + (\tau_K^{\max}h)^{1/2}\|\tau^{1/2}(\mathfrak{U}^{\text{DG}}\lambda - \lambda)\|_{\partial K}) \\
&\leq Ch^{-1}\|\lambda\|_{h,K}(\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_K + \|(\mathfrak{U}^{\text{DG}}\lambda - \lambda)\|_{\tau,\partial K}),
\end{aligned}$$

provided $\tau_K^{\max}h \leq C$. Therefore

$$\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|_K + \|\mathfrak{U}^{\text{DG}}\lambda - \lambda\|_{\tau,\partial K} \leq Ch^{-2}\|\lambda\|_{h,K}^2.$$

Summing over all K, we obtain the upper bound.

4.2 Estimates for HIP

Not surprisingly, the HIP method also has a similar estimate for the condition number.

Theorem 4.13. *Suppose $\mathcal{T}_h$ is a quasiuniform mesh of mesh size h. Assume there is a fixed constant $C_1, C_2 > 0$ such that*

$$\tau_K^{\min}h_K \geq C_1 \quad \textit{and} \quad \tau_K^{\max}h_K \leq C_2$$

for all elements K in all the meshes under consideration, where C_1 needs to be sufficiently large. Then, there are positive constants C_3 and C_4 independent of h such that

$$C_3\|\lambda\|_h^2 \le a_h^{\mathrm{IP}}(\lambda,\lambda) \le C_4 h^{-2}\|\lambda\|_h^2 \quad \text{for all } \lambda \in M_h. \tag{4.29}$$

As a result, the spectral condition number of A^{IP} satisfies

$$\kappa(A^{\mathrm{IP}}) \le Ch^{-2}.$$

Proof. We shall use the following identities [19]:

$$a_h^{\mathrm{IP}}(\lambda,\lambda) = \sum_K (\|\vec{\mathcal{Q}}^{\mathrm{DG}}\lambda\|_K^2 - \|\vec{\mathcal{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2 + \|\mathcal{U}^{\mathrm{DG}}\lambda - \lambda\|_{\tau,\partial K}^2) \tag{4.30}$$

$$= \sum_K (\|\vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2 - 2(\mathcal{U}^{\mathrm{DG}}\lambda - \lambda, \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\cdot\vec{n})_{\partial K} + \|\mathcal{U}^{\mathrm{DG}}\lambda - \lambda\|_{\tau,\partial K}^2) \tag{4.31}$$

$$= \sum_K ((\lambda, \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\cdot\vec{n})_{\partial K} + (\lambda, \tau_K(\mathcal{U}^{\mathrm{DG}}\lambda - \lambda))_{\partial K}). \tag{4.32}$$

First let us prove the lower bound. By Lemma 4.9 and 4.10, we just need to show $\|\vec{\mathcal{Q}}^{\mathrm{DG}}\lambda\|_K^2 \le a_h^{\mathrm{IP}}(\lambda,\lambda)$. From (4.30), it is clear that it suffices to prove

$$\|\vec{\mathcal{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2 \le \|\mathcal{U}^{\mathrm{DG}}\lambda - \lambda\|_{\tau,\partial K}^2. \tag{4.33}$$

To this end, starting with the following identity [19, Lemma 2.2]

$$(\vec{\mathcal{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda, \vec{v})_K = (\mathcal{U}^{\mathrm{DG}}\lambda - \lambda, \vec{v}\cdot\vec{n})_{\partial K}, \quad \forall \vec{v} \in V(K), \tag{4.34}$$

and a local scaling argument, we have

$$\|\vec{\mathcal{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2 \le C(\tau_K^{\min}h_K)^{-1/2}\|\mathcal{U}^{\mathrm{DG}}\lambda - \lambda\|_{\tau,\partial K}\|\vec{\mathcal{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K,$$

so (4.33) follows provided $\tau_K^{\min}h_K \ge 1/C^2$.

Next we prove the upper bound estimate. By (4.32),

$$
\begin{aligned}
a_h^{\mathrm{IP}}(\lambda,\lambda) &= \sum_K (\lambda, \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\cdot\vec{n})_{\partial K} + (\lambda, \tau_K(\mathcal{U}^{\mathrm{DG}}\lambda-\lambda))_{\partial K}) \\
&\le Ch^{-1}\sum_K (\|\lambda\|_{\partial K} h_K^{1/2})(\|\vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2 + \tau_K^{\max} h_K \|\mathcal{U}^{\mathrm{DG}}\lambda-\lambda\|_{\tau,\partial K}^2)^{1/2} \\
&\le Ch^{-1}\left(\sum_K \|\lambda\|_{\partial K}^2 h_K\right)^{1/2}\left(\sum_K (\|\vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2 + \|\mathcal{U}^{\mathrm{DG}}\lambda-\lambda\|_{\tau,\partial K}^2)\right)^{1/2},
\end{aligned}
$$

where in the last step we use the assumption $\tau_K^{\max} h_K \le C$. Hence it suffices to show that

$$
\sum_K (\|\vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2 + \|\mathcal{U}^{\mathrm{DG}}\lambda-\lambda\|_{\tau,\partial K}^2) \le C a_h^{\mathrm{IP}}(\lambda,\lambda). \tag{4.35}
$$

For this we use (4.31) and

$$
\begin{aligned}
|(\mathcal{U}^{\mathrm{DG}}\lambda-\lambda, \vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\cdot\vec{n})_{\partial K}| &\le C(\tau_K^{\min} h_K)^{-1/2}\|\mathcal{U}^{\mathrm{DG}}\lambda-\lambda\|_{\tau,\partial K}\|\vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K \\
&\le \hat{C}(\|\mathcal{U}^{\mathrm{DG}}\lambda-\lambda\|_{\tau,\partial K}^2 + \|\vec{\nabla}\,\mathcal{U}^{\mathrm{DG}}\lambda\|_K^2),
\end{aligned}
$$

with $\hat{C} = \frac{C}{2}(\tau_K^{\min} h_K)^{-1/2}$. Choosing $\tau_K^{\min} h_K$ sufficiently large, we can make $\hat{C}$ as small as we wish. In particular, applying this estimate with $2\hat{C} \le \gamma < 1$, in the identity (4.31), we see that (4.35) follows. Thus the proof is complete.

Remark 4.14. In the proof of Theorem 4.13, we put strong restrictions on the choices of τ to favor the arguments. But we suspect that this estimate (4.29) may still hold for less restrictive τ. That conjecture will be left for future investigation.

4.3 Numerical Experiments

In this section, we design and perform numerical experiments to investigate the decay rates of the errors in the HDG method. We will compare their performance with the continuous Galerkin method introduced in § 2.1. We will also investigate the dependence of the errors on the penalty parameter τ. We implemented the HDG method and performed experiments with the same meshes and domain as in § 3.3. The exact test solution is again $u(x,y) = \sin(x)e^{y/2}$.

We report three different discretization errors of the HDG method for $d = 0$ and $d = 1$ in Table 4-1 to 4-6, namely, $\|u - u_h\|_{L^2}$, $\|\vec{q} - \vec{q}_h\|_{L^2}$ and $\|Pu - \lambda_h\|_a$, where Pu is the L^2-orthogonal projection onto M_h. As seen from those tables, for fixed τ, all discretization errors converge in $O(h)$ for $d = 0$, and $O(h^2)$ for $d = 1$. These agree with the known theoretical estimateson HDG errors [20, Theorem 2.1] as well as the theorem we proved earlier, namely Theorem 4.5. Notice how the rate of convergence slows down if τ is set to $1/h$. This has been a traditional choice in many DG methods. However, for HDG methods, it is important that the penalty not be set to $1/h$ as seen from the last row of Tables 4-1 through 4-6.

Table 4-1. Discretization error $\|Pu - \lambda_h\|_a$ of HDG($d = 0$) for varying τ and h.

$\tau \diagdown h$	1	1/2	1/4	1/8	1/16	1/32
1/2	0.90388	0.00599	0.00284	0.00140	0.00069	0.00034
1	0.87065	0.01186	0.00566	0.00279	0.00139	0.00069
2	0.81477	0.02319	0.01119	0.00555	0.00277	0.00139
4	0.73273	0.04449	0.02191	0.01098	0.00552	0.00277
8	0.63661	0.07042	0.03853	0.02053	0.01067	0.00544
$\|e\|$	0.91435	0.00181	0.00043	0.00010	2.6e-05	6.6e-06
$1/\|e\|$	0.77696	0.07769	0.07219	0.07033	0.06975	0.06958

Table 4-2. Discretization error $\|u - u_h\|_{L^2}$ of HDG($d = 0$) for varying τ and h.

$\tau \diagdown h$	1	1/2	1/4	1/8	1/16	1/32
1/2	0.12823	0.03482	0.01741	0.00871	0.00435	0.00218
1	0.12115	0.03482	0.01741	0.00871	0.00436	0.00218
2	0.11020	0.03484	0.01743	0.00872	0.00436	0.00218
4	0.09626	0.03489	0.01747	0.00875	0.00437	0.00219
8	0.08305	0.03506	0.01762	0.00885	0.00443	0.00222
$\|e\|$	0.14014	0.03551	0.01776	0.00888	0.00444	0.00222
$1/\|e\|$	0.09623	0.03489	0.01798	0.00998	0.00661	0.00545

Next we test and compare the performance among the continuous Galerkin (CG) method, the HRT method and the HDG method with different τ. The comparisons are based on three kinds of different discretization errors, namely, the primal solution error $\|u - u_h\|_{L^2}$, the flux solution error $\|\vec{q} - \vec{q}_h\|_{L^2}$, and the *trace* solution error $\|\lambda - \lambda_h\|_a$.

Table 4-3. Discretization error $\|\vec{q}-\vec{q}_h\|_{L^2}$ of HDG($d=0$) for varying τ and h.

$\tau \diagdown h$	1	1/2	1/4	1/8	1/16	1/32
1/2	1.21571	0.00593	0.00283	0.00139	0.00069	0.00035
1	1.12791	0.01159	0.00559	0.00278	0.00139	0.00069
2	0.98672	0.02216	0.01094	0.00549	0.00276	0.00139
4	0.79281	0.04078	0.02093	0.01073	0.00546	0.00276
8	0.58004	0.07041	0.03854	0.02054	0.01067	0.00545
$\|e\|$	1.24191	0.00181	0.00043	0.00011	2.6e-05	6.6e-06
$1/\|e\|$	0.90225	0.06789	0.06318	0.06158	0.06107	0.06092

Table 4-4. Discretization error $\|Pu-\lambda_h\|_a$ of HDG($d=1$) for varying τ and h.

$\tau \diagdown h$	1	1/2	1/4	1/8	1/16	1/32
1/2	0.00170	0.00047	0.00012	3.2e-05	8.0e-06	2.0e-06
1	0.00214	0.00058	0.00015	3.8e-05	9.5e-06	2.4e-06
2	0.00332	0.00087	0.00022	5.6e-05	1.4e-05	3.5e-06
4	0.00588	0.00153	0.00039	9.8e-05	2.5e-05	6.2e-06
8	0.01084	0.00289	0.00074	0.00019	4.7e-05	1.2e-05
$\|e\|$	0.00167	0.00047	0.00012	3.2e-05	7.9e-06	2.0e-06
$1/\|e\|$	0.00486	0.00233	0.00114	0.00057	0.00028	0.00014

Table 4-5. Discretization error $\|u-u_h\|_{L^2}$ of HDG($d=1$) for varying τ and h.

$\tau \diagdown h$	1	1/2	1/4	1/8	1/16	1/32
1/2	0.00320	0.00079	0.00019	4.9e-05	1.2e-05	3.1e-06
1	0.00264	0.00066	0.00016	4.1e-05	1.0e-05	2.6e-06
2	0.00247	0.00062	0.00016	3.9e-05	9.7e-06	2.4e-06
4	0.00243	0.00061	0.00015	3.8e-05	9.5e-06	2.4e-06
8	0.00246	0.00061	0.00015	3.8e-05	9.5e-06	2.4e-06
$\|e\|$	0.00382	0.00155	0.00072	0.00035	0.00018	8.8e-05
$1/\|e\|$	0.00244	0.00061	0.00015	3.8e-05	9.5e-06	2.4e-06

Table 4-6. Discretization error $\|\vec{q}-\vec{q}_h\|_{L^2}$ of HDG($d=1$) for varying τ and h.

$\tau \diagdown h$	1	1/2	1/4	1/8	1/16	1/32
1/2	0.00454	0.00115	0.00029	7.3e-05	1.8e-05	4.6e-06
1	0.00472	0.00120	0.00030	7.7e-05	1.9e-05	4.8e-06
2	0.00538	0.00139	0.00035	8.9e-05	2.2e-05	5.6e-06
4	0.00725	0.00192	0.00049	0.00013	3.2e-05	7.9e-06
8	0.01127	0.00315	0.00084	0.00022	5.5e-05	1.4e-05
$\|e\|$	0.00455	0.00116	0.00029	7.3e-05	1.8e-05	4.6e-06
$1/\|e\|$	0.00649	0.00266	0.00124	0.00061	0.00030	0.00015

We first compare $\|u - u_h\|_{L^2}$ between the CG method (d=1) and the HRT method (d=1) in Figure 4-1. $x-$axis represents the logarithm of the degrees of freedom, and $y-$axis represents $\log(\|u - u_h\|_{L^2})$. The CG method (d=1) uses the finite element space consisting of *continuous* piecewise linear functions to approximate the exact solution. In contrast, The HDG method (d=1) uses the finite element space of *discontinuous* piecewise linear functions. We list three cases corresponding different τ (1/10000, 1 and 10000) for the latter. The results show that the CG method yields the best approximations for the primal variable u. However, the CG method does not yield good fluxes (which is one the primary motivations for constructing mixed and hybrid DG methods based on the dual form).

Next, we compare $\|\vec{q} - \vec{q}_h\|_{L^2}$ among the HRT method (d=0) and the HDG method (d=1) in Figure 4-2. Both methods use linear functions to approximate the flux solution $\vec{q}$, and the difference is that the latter uses the full space of linear piecewise functions, while the former uses only a subspace of it (the normal components need to be continuous across the interior faces). We see that while the performance of the HDG method remains relatively unaltered as τ goes from unit size to zero, it deteriorates if τ approaches infinity.

We proceed to compare $\|\lambda - \lambda_h\|_a$ among the HRT method (d=0) and the HDG method (d=0) in Figure 4-3. Both methods use the space of piecewise constant functions to approximate the trace solution. The HDG methods seem to have larger errors on coarser meshes, but their performance quickly become comparable to the HRT method when going to finer meshes.

In all these graphs, it may seem that the smaller τ is, the better the approximation gets, which confirms the results as indicated in Table 4-4. Moreover, the bigger τ is, the closer the HDG curve gets to the CG curve. As a matter of fact, the CG method can be considered as a limiting case of the HDG method by letting τ go to infinity [19].

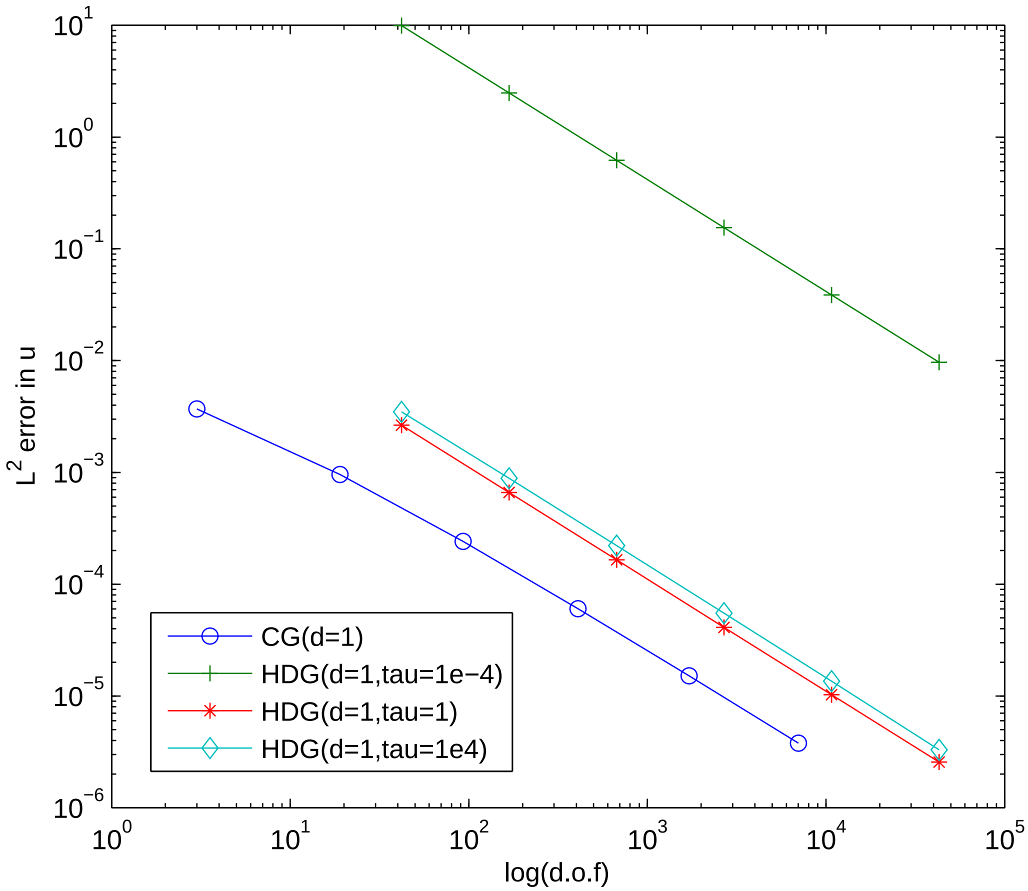

Figure 4-1. Comparison of $\log \|u - u_h\|_{L^2}$ among the CG method (d=1), the HRT method (d=0) and the HDG method (d=1) with different τ

.

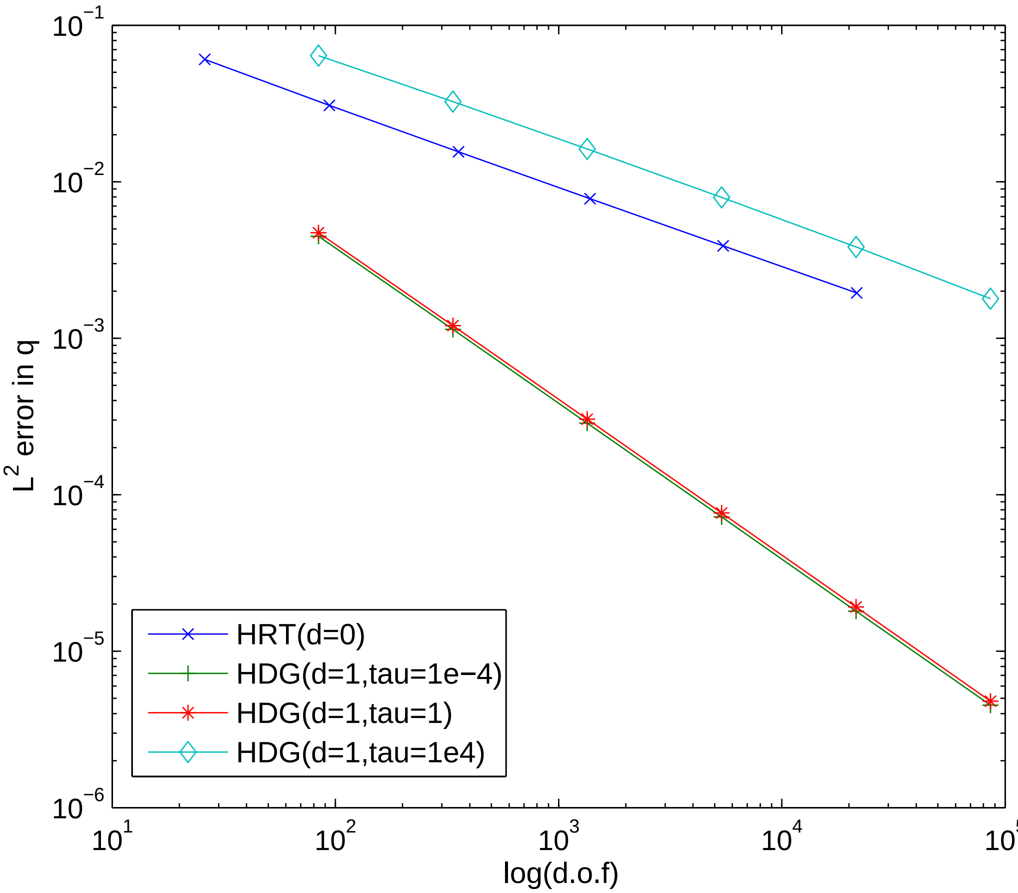

Figure 4-2. Comparison of $\log \|\vec{q} - \vec{q}_h\|_{L^2}$ among the the HRT method (d=0) and the HDG method (d=1) with different τ

.

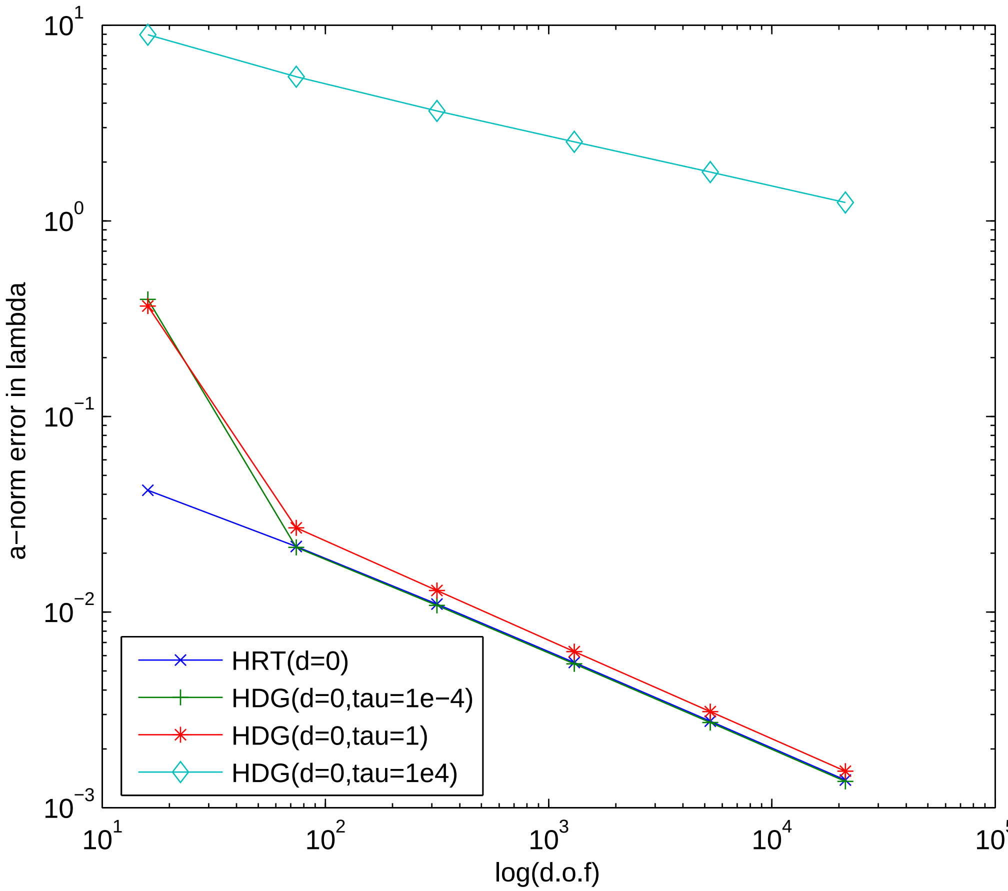

Figure 4-3. Comparison of $\log \|Pu - \lambda_h\|_a$ among the the HRT method (d=0) and the HDG method (d=0) with different τ

.

CHAPTER 5
MULTIGRID FOR HDG

The main contribution of this chapter is a provably efficient multigrid algorithm for the HDG method. While the algorithm itself is a straightforward generalization of what we presented in Chapter 3 for the HRT method, its analysis is more involved due to complexities (such as τ-dependence) in the estimates for the HDG method. Furthermore, we relax the regularity assumptions on the original boundary value problem in this chapter, resulting in a more generally applicable convergence theorem, but necessitating more technical proofs. The results of this chapter will be published in [16].

We use the same algorithm setting as in the HRT case, specifically we continue to use Algorithm 3.2. Our main goal is to prove a convergence result for this algorithm applied to the HDG case. Again, we use the abstract multigrid theory in [10], so that the convergence proof reduces to the verification of Condition 3.4 through Condition 3.6. Note that the proof of Condition 3.6 remains the same for HRT and HDG (smoothing subspaces are identical), so it leaves us only the first 2 conditions to verify.

Recalling the setting of Algorithm 3.2, note that there are many similarities when it is adapted to the HDG setting. As in § 3.1, we require a multilevel hierarchy of meshes and spaces and that the mesh $\mathscr{T}_h$ in which the solution is sought, is obtained by successive refinements of a coarse mesh $\mathscr{T}_1$ as in § 3.1. The HDG multilevel spaces M_k are identical to those in the HRT setting. The main difference is that the bilinear form $a_k(\cdot,\cdot)$ and the resulting matrices (and operators A_k) are now obtained from the HDG method. The smoothers depend on A_k, so they must also be modified. The intergrid transfer operator I_k remains the same.

The following is the main result of this chapter. It is proved under the following assumptions.

Assumption 5.1. Assume the following:

(i) The boundary value problem (2.12) admits the following regularity estimate for its solution:

$$\|\vec{q}\|_{H^s(\Omega)^n} + \|u\|_{H^{1+s}(\Omega)} \leq C\|f\|_{H^{-1+s}(\Omega)} \tag{5.1}$$

for some number $1/2 < s \leq 1$.

(ii) The assumption on the number of smoothings in Theorem 3.3 hold.

(iii) Assume that there is a fixed constant C such that

$$\tau_K^{\max} h_K \leq C$$

for all elements K in all meshes under consideration.

Note that in Theorem 3.3 we assumed the convexity of the domain Ω so that we have access to a full H^2-regularity result for u. The full regularity estimate, namely (5.1) with $s = 1$, is well known to hold when Ω is a convex domain in two or three dimensions [26]. Now we only assume (5.1) instead, which holds more generally. If Ω is a polygon (with no slits) then it is well known [22, 29] that the estimate (5.1) holds with $s > 1/2$. Note also that once (5.1) holds with $s > 1/2$, we can apply the projection $\mathit{\Pi}_h$ to $(\vec{q}, u)$.

Theorem 5.2. *Consider Algorithm 3.2 modified to apply to the HDG method as described above. Suppose Assumption 5.1 holds. Assume also that $d \geq 1$. Then there exists a positive $\delta < 1$, independent of the mesh size h_J, such that the error reducing operator $\mathcal{E}_J^{\mathrm{DG}}$ of the algorithm satisfies*

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

The following section is devoted to the proof of this theorem.

Remark 5.3. Numerical experiments indicate that the result even with zero degree (constant) $d = 0$ approximations, the conclusion of Theorem 5.2 may still hold. But the theoretical proof requires some tedious modifications, which are postponed for future research. Throughout the rest of this chapter, we always assume $d > 0$.

As in the HRT case, we may use the multigrid algorithm as either a linear iteration or as a preconditioner. For further details on how, see § 3.1.

5.1 Multigrid Convergence Analysis

We start with a lemma that reveals an equivalent norm generated by the HDG bilinear form.

Lemma 5.4. *For all $\lambda \in M_h$,*

$$C_1 \|\!|\lambda|\!\|_h \leq a_h(\lambda, \lambda) \leq C_2 \gamma_{h\tau}^{(1)} \|\!|\lambda|\!\|_h.$$

where $\gamma_{h\tau}^{(1)} = \max\{1 + \tau_K^{\max} h_K : K \in \mathscr{T}_h\}$.

Proof. The first inequality follows from Lemma 4.10, so it only remains to prove the upper bound. For this, note that

$$\begin{aligned} a_h(\lambda, \lambda) &= \|\vec{\mathfrak{Q}}^{\text{DG}}\lambda\|^2 + \|\mathfrak{U}^{\text{DG}}\lambda - \lambda\|_\tau^2 \\ &\leq 2\|\vec{\mathfrak{Q}}^{\text{DG}}\lambda - \vec{\mathfrak{Q}}^{\text{RT}}\lambda\|^2 + 2\|\vec{\mathfrak{Q}}^{\text{RT}}\lambda\|^2 + C(1 + \tau_K^{\max} h)\|\vec{\mathfrak{Q}}^{\text{RT}}\lambda\|^2 && \text{by (4.12)} \\ &\leq C\gamma_{h\tau}^{(1)} \|\vec{\mathfrak{Q}}^{\text{RT}}\lambda\|^2 && \text{by (3.24).} \end{aligned}$$

Hence the upper bound follows from the inequality

$$\|\vec{\mathfrak{Q}}^{\text{RT}}\lambda\| \leq C \|\!|\lambda|\!\|_h$$

proved in [24].

Lemma 5.5. *The following identities hold for all λ, η in M_J and all w in M_{J-1}:*

$$\vec{\mathfrak{Q}}^{\text{DG}}(I_J w) = -\vec{\nabla}\, w, \qquad (\forall\, d), \tag{5.2}$$

$$a_J(I_J w, \eta) = -(\vec{\nabla}\, w, \vec{\mathfrak{Q}}^{\text{DG}}\eta), \qquad \textit{if } d > 0, \tag{5.3}$$

$$(\vec{\mathfrak{Q}}^{\text{DG}}\lambda + \vec{\nabla}\, P_{J-1}\lambda, \vec{\nabla}\, w) = 0, \qquad \textit{if } d > 0, \tag{5.4}$$

$$a_J(\lambda - I_J P_{J-1}\lambda, \lambda) = \|\vec{\mathfrak{Q}}^{\text{DG}}\lambda + \vec{\nabla}\, P_{J-1}\lambda\|_c^2 + \|\mathfrak{U}^{\text{DG}}\lambda - \lambda\|_\tau^2, \qquad \textit{if } d > 0. \tag{5.5}$$

Proof. If $d > 0$, then the first identity follows from Lemma 4.4 (ii). If $d = 0$, it follows from the definition of $\vec{\mathfrak{Q}}^{\text{DG}}(I_J w)$, namely (4.4a), which reduces to

$$(\vec{\mathfrak{Q}}^{\text{DG}}(I_J w), \vec{r})_K - 0 = -\langle I_J w, \vec{r} \cdot \vec{n} \rangle_{\partial K} = -\langle w, \vec{r} \cdot \vec{n} \rangle_{\partial K} = -(\vec{\nabla} w, \vec{r})_K$$

for all constant vectors $\vec{r}$. Since $\vec{\mathfrak{Q}}^{\text{DG}}(I_J w)$ and $\vec{\nabla} v$ are constant vectors, (5.2) follows.

To prove (5.3), we use (5.2) in the definition of $a_J(\cdot, \cdot)$ to get

$$a_J(I_J w, \eta) = (-\vec{\nabla} w, \vec{\mathfrak{Q}}^{\text{DG}}\eta)_{\Omega_h} + \langle \tau(\mathfrak{U}^{\text{DG}}(I_J w) - w), \mathfrak{U}^{\text{DG}}\eta - \eta \rangle_{\partial\Omega_h} \tag{5.6}$$

and observe that the last term vanishes because of Lemma 4.4 (ii).

To prove (5.4), we again use (5.2):

$$\begin{aligned}
(\vec{\mathfrak{Q}}^{\text{DG}}\lambda + \vec{\nabla} P_{J-1}\lambda, \vec{\nabla} w) &= (\vec{\mathfrak{Q}}^{\text{DG}}\lambda, \vec{\nabla} w) + (\vec{\nabla} P_{J-1}\lambda, \vec{\nabla} w) \\
&= -(\vec{\mathfrak{Q}}^{\text{DG}}\lambda, \vec{\mathfrak{Q}}^{\text{DG}}(I_J w)) + a_{J-1}(P_{J-1}\lambda, w) \\
&= -(\vec{\mathfrak{Q}}^{\text{DG}}\lambda, \vec{\mathfrak{Q}}^{\text{DG}}(I_J w)) + a_J(\lambda, I_J w) = \langle \tau(\mathfrak{U}^{\text{DG}}\lambda - \lambda), \mathfrak{U}^{\text{DG}}(I_J w) - w \rangle.
\end{aligned}$$

which holds for all $d > 0$. The last term then vanishes because of Lemma 4.4 (ii) and we get (5.4).

Finally, to prove (5.5),

$$\begin{aligned}
a_J(\lambda - I_J P_{J-1}\lambda, \lambda) &= a_J(\lambda, \lambda) - a_{J-1}(P_{J-1}\lambda, P_{J-1}\lambda) \\
&= (\vec{\mathfrak{Q}}^{\text{DG}}\lambda, \vec{\mathfrak{Q}}^{\text{DG}}\lambda) - (\vec{\nabla} P_{J-1}\lambda, \vec{\nabla} P_{J-1}\lambda) + \|\mathfrak{U}^{\text{DG}}\lambda - \lambda\|_\tau^2 \\
&= ((\vec{\mathfrak{Q}}^{\text{DG}}\lambda + \vec{\nabla} P_{J-1}\lambda), (\vec{\mathfrak{Q}}^{\text{DG}}\lambda - \vec{\nabla} P_{J-1}\lambda)) + \|\mathfrak{U}^{\text{DG}}\lambda - \lambda\|_\tau^2.
\end{aligned}$$

Then it suffices to note that the term $\vec{\mathfrak{Q}}^{\text{DG}}\lambda - \vec{\nabla} P_{J-1}\lambda$ can be replaced by $\vec{\mathfrak{Q}}^{\text{DG}}\lambda + \vec{\nabla} P_{J-1}\lambda$ due to (5.4) whenever $d > 0$.

5.1.1 Norm of Prolongation Operator(Condition 3.4)

In this subsection we prove the following theorem.

Theorem 5.6. *For all v in M_{k-1}, we have*

$$a_k(I_k v, I_k v) = a_{k-1}(v,v), \qquad \text{if } d > 0.$$

Lemma 5.7. *If $d = 0$, then for all w in $P_1(K)$,*

$$\|\mathcal{U}^{\mathrm{DG}}(I_J w) - w\|_{\tau,\partial K} \le C(\tau_K^{\max} h_K)^{1/2} \|\vec{\nabla} w\|_K, \tag{5.7}$$

$$\|\mathcal{U}^{\mathrm{DG}}(I_J w) - I_J w\|_{\tau,\partial K} \le C(\tau_K^{\max} h_K)^{1/2} \|\vec{\nabla} w\|_K. \tag{5.8}$$

$$\|\mathcal{U}^{\mathrm{DG}}(I_J w) - w\|_K \le C h_K \|\vec{\nabla} w\|_K. \tag{5.9}$$

Proof. Because of (2.13),

$$\langle \tau(\mathcal{U}^{\mathrm{DG}}(I_J w) - w), \mathcal{U}^{\mathrm{DG}}(I_J w) - w_0\rangle_{\partial K} = 0.$$

for any constant w_0. Hence,

$$\|\mathcal{U}^{\mathrm{DG}}(I_J w) - w\|_{\tau,\partial K} \le \|w - w_0\|_{\tau,\partial K}. \tag{5.10}$$

Applying a local trace inequality,

$$\|w - w_0\|_{\tau,\partial K} \le \left(\frac{\tau_K^{\max}}{h_K}\right)^{1/2} \|w - w_0\|_K \le C(\tau_K^{\max} h_K)^{1/2} \|\vec{\nabla} w\|_K, \tag{5.11}$$

where we have used Friedrichs inequality after choosing w_0 to be the mean of w on K. This proves (5.7).

Inequality (5.8) follows from (5.7) by triangle inequality once we prove that

$$\|I_J w - w\|_{\tau,\partial K} \le C(\tau_K^{\max} h_K)^{1/2} \|\vec{\nabla} w\|_K. \tag{5.12}$$

Since the mean of $I_J w - w$ vanishes on each face F of ∂K, applying Friedrichs inequality, we have

$$\|I_J w - w\|_F \le C h_K \|(\vec{\nabla} w)_F\|_F, \tag{5.13}$$

where $(\vec{\nabla}\, w)_F$ denotes the tangential gradient. More generally, let us denote by $\vec{\kappa}_F$ the tangential component on F of any constant vector $\vec{\kappa}$. Then it is easy to see that

$$h_K^{1/2}\|\vec{\kappa}_F\| \le C\|\vec{\kappa}\|_K.$$

Using this with $\vec{\kappa} = \vec{\nabla}\, w$, we see that (5.13) implies (5.12).

For (5.9), we use a standard local estimate for linear functions,

$$C\|\mathfrak{U}^{\rm DG}(I_J w) - w\|_K \le h_K\|\,\vec{\nabla}(\mathfrak{U}^{\rm DG}(I_J w) - w)\|_K + h_K^{1/2}\|\mathfrak{U}^{\rm DG}(I_J w) - w)\|_F$$

for any face F of K. Choosing $F = F_{\max}$, a face where τ assumes its maximum value,

$$\begin{aligned}
C\|\mathfrak{U}^{\rm DG}(I_J w) - w\|_K &\le h_K\|\,\vec{\nabla}\, w\|_K + h_K^{1/2}(\tau_K^{\max})^{-1/2}\|\mathfrak{U}^{\rm DG}(I_J w) - w\|_{\tau,F_{\max}}\\
&\le h_K\|\,\vec{\nabla}\, w\|_K + h_K^{1/2}(\tau_K^{\max})^{-1/2}\|w - w_0\|_{\tau,\partial K}\\
&\le h_K\|\,\vec{\nabla}\, w\|_K + h_K^{1/2}\|w - w_0\|_{\partial K},
\end{aligned}$$

where we have used (5.10). Then we finish the proof as in (5.11).

Proof of Theorem 5.6. Obviously, it suffices to prove the inequalities for $k = J$. To prove the $d > 0$ case, we use two identities of Lemma 5.5:

$$\begin{aligned}
a_J(I_J v, I_J v) &= -(\vec{\nabla}\, v, \vec{\mathfrak{Q}}^{\rm DG}(I_J v)) && \text{by (5.3)}\\
&= (a\,\vec{\nabla}\, v, \vec{\nabla}\, v) && \text{by (5.2)}.
\end{aligned}$$

To prove the next inequality for the $d = 0$ case, we again use (5.2) of Lemma 5.5, hence we can put $\vec{\mathfrak{Q}}^{\rm DG}(I_J v) = -\,\vec{\nabla}\, v$ in the definition of $a_j(\cdot,\cdot)$ to get

$$\begin{aligned}
a_J(I_J v, I_J v) &= (\vec{\nabla}\, v, \vec{\nabla}\, v) + \|\mathfrak{U}^{\rm DG}(I_J v) - I_J v\|_\tau^2\\
&\le a_{J-1}(v, v) + C\tau_K^{\max} h_J\|\,\vec{\nabla}\, v\|^2,
\end{aligned}$$

where the last step was due to (5.8) of Lemma 5.7. This proves the theorem.

5.1.2 Regularity and Approximation Property (Condition 3.5)

We begin with a simple consequence of Theorem 5.6.

Lemma 5.8. *For all μ in M_J, we have*

$$a_{J-1}(P_{J-1}\mu, P_{J-1}\mu) \leq \|\mu\|_a^2, \qquad \text{if } d > 0.$$

Proof. The estimate follows from

$$\begin{aligned}(\vec{\nabla} P_{J-1}\mu, \vec{\nabla} P_{J-1}\mu) &= a_{J-1}(P_{J-1}\mu, P_{J-1}\mu) = a_J(\mu, I_J P_{J-1}\mu) \\ &\leq \|\mu\|_a \, \|I_J P_{J-1}\mu\|_a,\end{aligned}$$

and applying Theorem 5.6 to the last term.

Next we define a new operator, S, defined on each edge of the mesh. Let λ be the restriction of a function in M_h on ∂K for some mesh element K. Let F_i denote the face of K opposite to the ith vertex of K. Then define $S_i^K\lambda$ in $P_{k+1}(K)$ by

$$\langle S_i^K\lambda, \eta\rangle_{F_i} = \langle \lambda, \eta\rangle_{F_i} \qquad \text{for all } \eta \in P_{k+1}(F_i), \tag{5.14a}$$

$$(S_i^K\lambda, v)_K = (\mathcal{U}^{\text{DG}}\lambda, v)_K \qquad \text{for all } v \in P_k(K), \tag{5.14b}$$

and, considering all the $n+1$ faces of K, define

$$(\lambda, \mu)_S = \sum_{K\in\mathscr{T}_h} \frac{1}{n+1} \sum_{i=1}^{n+1} (S_i^K\lambda, S_i^K\mu)_K \qquad \text{and} \qquad \|\lambda\|_S^2 = (\lambda, \lambda)_S.$$

Lemma 5.9. *Equations (5.14a) and (5.14b) uniquely define a $S_i^K\lambda$ in $P_{k+1}(K)$. Furthermore, for all λ in M_h,*

$$\mathcal{U}^{\text{DG}}\lambda|_K = \Pi_{W_h}(S_i^K\lambda), \tag{5.15}$$

$$C_5\|\lambda\|_J \leq \|\lambda\|_S \leq C_6\gamma_{h\tau}^{(1/2)}\|\lambda\|_J, \tag{5.16}$$

$$\|\vec{\nabla}(S_i^K\lambda)\|_K \leq C\|\vec{\mathcal{Q}}^{\text{DG}}\lambda\|_K \tag{5.17}$$

where $\gamma_{h\tau}^{(1/2)} = \max\{1 + (\tau_K^{\max} h_K)^{1/2} : K \in \mathscr{T}_h\}$.

Proof. Since (5.14) forms a square system for $S_i\lambda$, to show that it has a unique solution, it suffices to show that the only solution when the right hand sides are zero is the trivial solution. That this is indeed the case is an immediate consequence of [20, Lemma 2.1]. The identity (5.15) is obvious from (5.14b). Let us prove the remaining assertions.

We prove (5.16) by a scaling argument. To this end, consider a fixed reference simplex $\hat{K}$, with an arbitrarily chosen face $\hat{F}$, and define $\psi_{\hat{\lambda},\hat{q}}$ by

$$\langle \psi_{\hat{\lambda},\hat{q}}, \eta\rangle_{\hat{F}} = \langle \hat{\lambda}, \eta\rangle_{\hat{F}} \qquad \text{for all } \eta \in P_{k+1}(\hat{F}),$$

$$(\psi_{\hat{\lambda},\hat{q}}, v)_{\hat{K}} = (\hat{q}, v)_{\hat{K}} \qquad \text{for all } v \in P_k(\hat{K}).$$

It is easy to see that the $\|\psi_{\hat{\lambda},\hat{q}}\|_{\hat{K}}$ and $(\|\hat{\lambda}\|_{\hat{F}}^2 + \|\hat{q}\|_{\hat{K}}^2)^{1/2}$ are equivalent norms. Mapping to any element K such that $\hat{F}$ gets mapped to the face F_i of K, and relating $\psi_{\hat{\lambda},\hat{q}}$ to $S_i^K\lambda$, we have

$$C_5(h_K\|\lambda\|_{F_i}^2 + \|\mathcal{U}^{\text{DG}}\lambda\|_K^2) \le \|S_i^K\lambda\|_K^2 \le C(h_K\|\lambda\|_{F_i}^2 + \|\mathcal{U}^{\text{DG}}\lambda\|_K^2)$$

Summing over all faces $F_i \subset \partial K$,

$$C_5(h_K\|\lambda\|_{\partial K}^2) \le \sum_{i=1}^{n+1} \|S_i^K\lambda\|_K^2 \le C\left(h_K\|\lambda\|_{\partial K}^2 + C(1+(\tau_K^{\max}h_K)^{1/2})^2 h_K\|\lambda\|_{\partial K}^2\right) \tag{5.18}$$

where we have used Theorem 4.1. Now, to obtain (5.16), we need only sum over all K.

To prove (5.17), first observe that if λ takes a constant value κ on the boundary of some mesh element ∂K, then $S_i^K\lambda \equiv \kappa$. This is because $\mathcal{U}^{\text{DG}}\lambda \equiv \kappa$ by Lemma 4.4(i), so the function κ satisfies both the equations of (5.14). Therefore, by the unique solvability of (5.14), $S_i^K\lambda \equiv \kappa$. A consequence of this fact is that for any λ, we have

$$\vec{\nabla}(S_i^K(m_K(\lambda)) = 0$$

where $m_K(\lambda)$ is as in (2.9). Therefore,

$$
\begin{aligned}
\| \vec{\nabla}(S_i^K \lambda)\|_{L^2(K)} &= \| \vec{\nabla}\, S_i^K(\lambda - m_K(\lambda))\|_{L^2(K)}, \\
&\le C h_K^{-1} \| S_i^K(\lambda - m_K(\lambda))\|_{L^2(K)} && \text{(by an inverse inequality)} \\
&\le C h_K^{-1}(1 + (\tau_K^{\max} h_K)^{1/2}) h_K^{1/2} \|\lambda - m_K(\lambda)\|_{L^2(\partial K)} && \text{(by (5.18))} \\
&\le C \gamma_{h\tau}^{(1/2)} |\!|\!| \lambda - m_K(\lambda) |\!|\!|_{h,K},
\end{aligned}
$$

so (5.17) follows from Lemma 4.10.

Next, we define a map $\lambda \longmapsto \tilde{\lambda}$ from M_J into M_J as follows. First, given λ in M_J, let ϕ_λ be the unique function in M_J satisfying

$$
(\phi_\lambda, \mu)_S = a_h(\lambda, \mu), \quad \forall \mu \in M_J. \tag{5.19}
$$

This equation is uniquely solvable for ϕ_λ in M_J, because if the right-hand side is zero, then by (5.16) of Lemma 5.9, we have that $\phi_\lambda = 0$. Next, let $f_\lambda = \mathfrak{U}^{\mathrm{DG}} \phi_\lambda$ and define $\tilde{\lambda} \in M_J$ to be the unique solution of the equation

$$
a_h(\tilde{\lambda}, \mu) = (f_\lambda, \mathfrak{U}^{\mathrm{DG}} \mu), \quad \forall \mu \in M_J. \tag{5.20}
$$

Lemma 5.10. *The following statements hold for all λ in M_J:*

$$
\|f_\lambda\| \le \|\phi_\lambda\|_S \le C \|A_J \lambda\|_J \tag{5.21}
$$

$$
\|\lambda - \tilde{\lambda}\|_a \le C h_J \|A_J \lambda\|_J \tag{5.22}
$$

$$
\|f_\lambda\|_{H^{-1}(\Omega)} \le C d_\tau \|\lambda\|_a. \tag{5.23}
$$

where $d_\tau = 1 + (1 - \delta_{d0})(\tau_K^{\max} h_J)^{1/2}$.

Proof. The proofs of (5.21) and (5.22) are similar to the proof of 5.10. The only difference is that we now use the estimates of Lemma 5.9. To prove (5.21), first observe that

$$
f_\lambda = \Pi_{W_h} S_i^K \phi_\lambda
$$

by (5.15) of Lemma 5.9. Therefore,

$$\|f_\lambda\|^2 = \sum_{K\in\mathscr{T}_h} \frac{1}{n+1}\sum_{i=1}^{n+1} \|\Pi_{W_h}(S_i^K\phi_\lambda)\|_K^2 \le \|\phi_\lambda\|_S^2.$$

which is the first of the inequalities in (5.21). Moreover,

$$\begin{aligned}
\|\phi_\lambda\|_S^2 &= \sup_{\mu\in M_J} \frac{(\phi_\lambda,\mu)_S}{\|\mu\|_S} = \sup_{\mu\in M_J}\frac{a_J(\lambda,\mu)}{\|\mu\|_S} && \text{by (5.19)},\\
&= \sup_{\mu\in M_J}\frac{(A_J\lambda,\mu)_J}{\|\mu\|_S} \le \sup_{\mu\in M_J}\frac{(A_J\lambda,\mu)_J}{C_5\|\mu\|_J} && \text{by (5.16) of Lemma 5.9},\\
&\le C\|A_J\lambda\|_J,
\end{aligned}$$

thus completing the proof of (5.21).

To prove (5.22), let us first note that we can rewrite (5.19) and (5.20) as follows:

$$\begin{aligned}
a(\lambda,\mu) &= \sum_{K\in\mathscr{T}_h}\frac{1}{n+1}\sum_{i=1}^{n+1}(S_i^K\phi_\lambda, S_i^K\mu)_K,\\
a(\tilde\lambda,\mu) &= \sum_{K\in\mathscr{T}_h}\frac{1}{n+1}\sum_{i=1}^{n+1}(\Pi_{W_h}S_i^K\phi_\lambda, S_i^K\mu)_K.
\end{aligned}$$

To get the last identity, we have again used Lemma 5.9, whereby $f_\lambda = \mathfrak{U}^{\mathrm{DG}}\phi_\lambda = \Pi_{W_h}(S_i^K\phi_\lambda)$ on any element K. Subtracting, and setting $\mu = \lambda-\tilde\lambda$, we get

$$\begin{aligned}
\|\lambda-\tilde\lambda\|_a^2 &= \sum_{K\in\mathscr{T}_h}\frac{1}{n+1}\sum_{i=1}^{n+1}((I-\Pi_{W_h})S_i^K\phi_\lambda, S_i^K(\lambda-\tilde\lambda))\\
&= \sum_{K\in\mathscr{T}_h}\frac{1}{n+1}\sum_{i=1}^{n+1}(S_i^K\phi_\lambda, (I-\Pi_{W_h})S_i^K(\lambda-\tilde\lambda)).
\end{aligned}$$

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

$$\begin{aligned}
\|\lambda-\tilde\lambda\|_a^2 &\le C\|\phi_\lambda\|_S\Big(\sum_{K\in\mathscr{T}_J}\frac{1}{n+1}\sum_{i=1}^{n+1}h_J^2|S_i^K(\lambda-\tilde\lambda)|_{H^1(K)}^2\Big)^{1/2}\\
&\le Ch_J\,\|\phi_\lambda\|_S\;\|\vec{\mathfrak{Q}}^{\mathrm{DG}}(\lambda-\tilde\lambda)\|\\
&\le Ch_J\,\|A_J\lambda\|_J\;\|\lambda-\tilde\lambda\|_a
\end{aligned}$$

by (5.21) and (5.17) of Lemma 5.9. Canceling the common factor above, we obtain (5.22).

Next, let us prove (5.23). To this end, given any ψ in $H_0^1(\Omega)$, let ψ_{J-1} in M_{J-1} denote a function satisfying

$$\|\vec{\nabla}\psi_{J-1}\| \le C\|\vec{\nabla}\psi\| \qquad \text{and} \qquad \|\psi-\psi_{J-1}\| \le Ch_J\|\vec{\nabla}\psi\|. \tag{5.24}$$

Such approximations are well known to exist [34]. Then,

$$\begin{aligned}\|f_\lambda\|_{H^{-1}(\Omega)} &= \sup_{\psi\in H_0^1(\Omega)} \frac{(f_\lambda,\psi)}{\|\vec{\nabla}\psi\|} \\ &= \sup_{\psi\in H_0^1(\Omega)} \frac{(f_\lambda,(\psi-\psi_{J-1})\,+\,(\psi_{J-1}-\mathfrak{U}^{\mathrm{DG}}(I_J\psi_{J-1}))\,+\,\mathfrak{U}^{\mathrm{DG}}(I_J\psi_{J-1}))}{\|\vec{\nabla}\psi\|}.\end{aligned}$$

Now, since

$$\begin{aligned}(f_\lambda,\psi-\psi_{J-1}) &\le C\|f_\lambda\|h_J\|\vec{\nabla}\psi\|, && \text{by (5.24)},\\ (f_\lambda,\psi_{J-1}-\mathfrak{U}^{\mathrm{DG}}(I_J\psi_{J-1})) &= 0, && \text{by Lemma 4.4(ii), if } d>0,\\ (f_\lambda,\mathfrak{U}^{\mathrm{DG}}(I_J\psi_{J-1})) &= a_J(\tilde{\lambda},I_J\psi_{J-1}) && \text{by (5.20)},\\ &\le \|\tilde{\lambda}\|_a C d_\tau\|\vec{\nabla}\psi_{J-1}\|, && \text{by Theorem 5.6},\end{aligned}$$

the terms in the supremum can be bounded accordingly to get that

$$\|f_\lambda\|_{H^{-1}(\Omega)} \le Cd_\tau\|\tilde{\lambda}\|_a + Ch_J\|f_\lambda\|.$$

Finally, since (5.21) implies that $\|f_\lambda\| \le C\|A_J\lambda\|_J$, and since (5.22) implies

$$\|\tilde{\lambda}\|_a \le \|\lambda\|_a + Ch_J\|A_J\lambda\|_J,$$

we get

$$\|f_\lambda\|_{H^{-1}(\Omega)} \le Cd_\tau\|\lambda\|_a + Ch_J\|A_J\lambda\|_J.$$

The last term can be estimated by

$$\begin{aligned}
\|A_J\lambda\|_J^2 &= (A_J\lambda, A_J\lambda)_J \\
&= (A_J A_J^{1/2}\lambda, A_J^{1/2}\lambda)_J \\
&\leq \lambda_{max}(A_J)(A_J^{1/2}\lambda, A_J^{1/2}\lambda)_J \\
&= \lambda_{max}(A_J)(A_J^{1/2}A_J^{1/2}\lambda, \lambda)_J \\
&= \lambda_{max}(A_J)(A_J\lambda, \lambda)_J \\
&= \lambda_{max}(A_J)a(\lambda, \lambda) \\
&\leq Ch_J^{-2}\|\lambda\|_a^2,
\end{aligned}$$

where the last inequality was due to Theorem 4.8. Hence we finish the proof.

Now, let $\tilde{u}$ be the unique function in $H_0^1(\Omega)$ that solves

$$(\vec{\nabla}\,\tilde{u}, \vec{\nabla}\,v) = (f_\lambda, v), \quad \forall v \in H_0^1(\Omega), \tag{5.25}$$

and let $\tilde{u}_{J-1}$ be the unique function in M_{J-1} satisfying

$$(\vec{\nabla}\,\tilde{u}_{J-1}, \vec{\nabla}\,v) = (f_\lambda, v), \quad \forall v \in M_{J-1}. \tag{5.26}$$

Lemma 5.11. *If $d > 0$, then $P_{J-1}\tilde{\lambda} - \tilde{u}_{J-1} = 0$.*

Proof. Observe that for all w in M_{J-1},

$$\begin{aligned}
(\vec{\nabla}\,P_{J-1}\tilde{\lambda}, \vec{\nabla}\,w) &= a_{J-1}(P_{J-1}\tilde{\lambda}, w) = a_J(\tilde{\lambda}, I_J w) \\
&= (f_\lambda, \mathcal{U}^{\mathrm{DG}}(I_J w)),
\end{aligned} \tag{5.27}$$

by (5.20). Now, if $d > 0$, by Lemma 4.4 (ii), we know that $\mathcal{U}^{\mathrm{DG}}(I_J w) - w = 0$. Hence we have

$$(\vec{\nabla}\,P_{J-1}\tilde{\lambda}, \vec{\nabla}\,w) = (f_\lambda, w) \qquad \forall w \in M_{J-1},$$

which is the same equation satisfied by $\tilde{u}_{J-1}$. Hence $P_{J-1}\tilde{\lambda}$ and $\tilde{u}_{J-1}$ coincide if $d > 0$.

Lemma 5.12. *If s is as in Assumption 5.1, for any λ in M_J,*

$$\|\vec{\mathbb{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla} P_{J-1}\lambda\|^2 \le Ch_J^2\|A_J\lambda\|_J^2 + C\gamma_\tau h^{2s} a_J(\lambda,\lambda)^{1-s}\|A_J\lambda\|_J^{2s}.$$

where $\gamma_\tau = (1+\tau_K^{\max 2})(1+(\tau_K^{\max}h)^{1-s})$.

Proof. First, let us split the term requiring estimation as

$$\vec{\mathbb{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla}(P_{J-1}\lambda) = \sum_{i=1}^{6} t_i,$$

where

$$\begin{aligned}
t_1 &= \vec{\mathbb{Q}}^{\mathrm{DG}}(\lambda - \tilde{\lambda}),\\
t_2 &= \vec{\mathbb{Q}}^{\mathrm{DG}}(\tilde{\lambda} - P_h^M\tilde{u}),\\
t_3 &= \vec{\mathbb{Q}}^{\mathrm{DG}}(P_h^M\tilde{u} - I_J\tilde{u}_{J-1}),\\
t_4 &= \vec{\mathbb{Q}}^{\mathrm{DG}}(I_J\tilde{u}_{J-1}) - (-\vec{\nabla}\tilde{u}_{J-1}),\\
t_5 &= \vec{\nabla} P_{J-1}\tilde{\lambda} - \vec{\nabla}\tilde{u}_{J-1},\\
t_6 &= \vec{\nabla} P_{J-1}(\lambda - \tilde{\lambda}).
\end{aligned}$$

These terms are bounded as follows:

$$\begin{aligned}
\|t_1\| &\le \|\lambda - \tilde{\lambda}\|_a \le Ch_J\|A_J\lambda\|_J && \text{by Lemma 5.10},\\
\|t_2\| &\le \|\tilde{\lambda} - P_h^M\tilde{u}\|_a \le \|\tilde{q} - \Pi_h^V\tilde{q}\|_c && \text{by (4.23) of Theorem 4.5},\\
&\le Ch^s(|\tilde{q}|_{H^s(\Omega)} + \tau_K^{\max}|\tilde{u}|_{H^s(\Omega)}) && \text{by (4.21)},
\end{aligned}$$

where $\tilde{q} = -\vec{\nabla}\tilde{u}$. For t_3, we use (4.5) of Theorem 4.1 to get that

$$\|\vec{\mathbb{Q}}^{\mathrm{DG}}(P_h^M\tilde{u} - \tilde{u}_{J-1})\|^2 \le C\gamma_{h\tau}^{(1)}h^{-2}\|P_h^M(\tilde{u} - \tilde{u}_{J-1})\|_h^2 \le C\gamma_{h\tau}^{(1)}h^{-2}\|\tilde{u} - \tilde{u}_{J-1}\|_h^2, \tag{5.28}$$

where $\gamma_{h\tau}^{(1)} = \max\{1 + h_K \tau_K^{\max} : K \in \mathscr{T}_h\}$. By a local trace inequality, we can estimate the mesh dependent norm above by interior norms as follows:

$$C\|\tilde{u} - \tilde{u}_{J-1}\|_{h,K}^2 \leq \|\tilde{u} - \tilde{u}_{J-1}\|_K^2 + h_K^2 \| \vec{\nabla}(\tilde{u} - \tilde{u}_{J-1})\|_K^2. \tag{5.29}$$

Since $\tilde{u}_{J-1}$ is a standard Galerkin approximation [14] of $\tilde{u}$, we have

$$\| \vec{\nabla}(\tilde{u} - \tilde{u}_{J-1})\| \leq Ch^s |\tilde{u}|_{H^{1+s}(\Omega)}. \tag{5.30}$$

Furthermore, a standard duality argument [14, 30] yields

$$\|\tilde{u} - \tilde{u}_{J-1}\| \leq Ch^s \| \vec{\nabla}(\tilde{u} - \tilde{u}_{J-1})\| \leq Ch^{1+s} |\tilde{u}|_{H^{1+s}(\Omega)}. \tag{5.31}$$

Summing (5.29) over all elements and using (5.30) and (5.31), we can estimate $\|\tilde{u} - \tilde{u}_{J-1}\|_h$. Returning to (5.28) and using this bound, we have

$$\|t_3\| \leq C(\gamma_{h\tau}^{(1)})^{1/2} h^s |u|_{H^{1+s}(\Omega)}.$$

Proceeding to the succeeding terms,

$$\begin{aligned}
&\|t_4\| = 0, && \text{by (5.2) of Lemma 5.5,}\\
&\|t_5\| \leq Ch_J \|A_J \lambda\|_J, \quad \text{if } d > 0, &&\\
&\|t_6\| \leq C\|\lambda - \tilde{\lambda}\|_a \leq Ch\|A_J\lambda\|_J, && \text{by Lemmas 5.8 and 5.10.}
\end{aligned}$$

Combining these estimates for all t_i, we obtain

$$\begin{aligned}
\|\vec{\mathbb{Q}}^{\mathrm{DG}}\lambda + \vec{\nabla} P_{J-1}\lambda\|^2 &\leq Ch_J^2 \|A_J\lambda\|_J^2 + C\gamma_{h\tau}^{(1)} h_J^{2s} |\tilde{u}|_{H^{1+s}(\Omega)}^2 + Ch^{2s}(|\tilde{q}|_{H^s(\Omega)} + \tau_K^{\max}|\tilde{u}|_{H^s(\Omega)})^2\\
&\leq Ch_J^2 \|A_J\lambda\|_J^2 + C(1 + \tau_K^{\max 2}) h^{2s} \|f_\lambda\|_{H^{-1+s}(\Omega)}^2
\end{aligned} \tag{5.32}$$

by the regularity assumption (5.1). Since $H^{-1+s}(\Omega)$ is an interpolation space [4] in the scale of intermediate spaces between $H^{-1}(\Omega)$ and $L^2(\Omega)$, we know that

$$\|f_\lambda\|_{H^{-1+s}(\Omega)}^2 \leq \|f_\lambda\|_{H^{-1}(\Omega)}^{2(1-s)} \|f_\lambda\|^{2s}.$$

Therefore,

$$\|f_\lambda\|_{H^{-1+s}(\Omega)}^2 \leq C d_\tau^{2(1-s)} a(\lambda,\lambda)^{1-s} \|A_J\lambda\|_J^{2s}$$

by (5.23) and (5.21) of Lemma 5.10. Note that $d_\tau^{2(1-s)} \leq C(1+(\tau_K^{\max} h)^{1-s})$. Returning to (5.32) and using the resulting bound, we then obtain the estimate of the lemma.

Theorem 5.13. *Assume $\tau_K^{\max} h < C$. Then there exist $0 < \alpha \leq 1$ and C_1 such that*

$$a_k((I - I_k P_{k-1})v, v) \leq C_1 \left(\frac{\|A_k v\|_k^2}{\lambda_k} \right)^\alpha a_k(v,v)^{1-\alpha}, \quad \forall v \in M_k,\ k = 1, \ldots, J,$$

Proof. The inequality is known [6, 10] to hold for all $k < J$. Hence it suffices to prove the inequality when $k = J$. By the identities (5.5) of Lemma 5.5, we know that

$$a_k((I - I_k P_{k-1})v, v) \leq \|\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}\, P_{J-1}\lambda\|_c^2 + \|\mathfrak{U}^{\rm DG}\lambda - \lambda\|_\tau^2, \qquad \text{if } d > 0.$$

Since the terms involving $\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}\, P_{J-1}\lambda$ can be bounded as in Lemma 5.12, let us first investigate the remaining term involving $\mathfrak{U}^{\rm DG}\lambda - \lambda$. To this end, the following inequality will be helpful:

$$\|\mathfrak{U}^{\rm DG}(\lambda - I_J P_{J-1}\lambda) - (\lambda - I_J P_{J-1}\lambda)\|_{\tau,\partial K} \leq C\sqrt{\tau_K^{\max} h_K}\, \|\vec{\mathfrak{Q}}^{\rm DG}(\lambda - I_J P_{J-1}\lambda)\|_K.$$

This is due to (4.12) of Lemma 4.4. By Lemma 3.7, we also know that

$$\vec{\mathfrak{Q}}^{\rm DG}(I_J P_{J-1}\lambda) = -\vec{\nabla}\, P_{J-1}\lambda.$$

Thus

$$\begin{aligned} \|\mathfrak{U}^{\rm DG}(\lambda - I_J P_{J-1}\lambda) - (\lambda - I_J P_{J-1}\lambda)\|_{\tau,\partial K} &\leq C\sqrt{\tau_K^{\max} h_K}\, \|\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}(P_{J-1}\lambda)\|_K \\ &\leq C\sqrt{\tau_K^{\max} h_K}\, \|\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}(P_{J-1}\lambda)\|_K, \qquad (5.33) \end{aligned}$$

where the last inequality was because $\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}(P_{J-1}\lambda) = J_K(\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}(P_{J-1}\lambda))$.

Now, if $d > 0$, then

$$\|\mathfrak{U}^{\rm DG}\lambda - \lambda\|_{\tau,\partial K} = \|\mathfrak{U}^{\rm DG}(\lambda - I_J P_{J-1}\lambda) - (\lambda - I_J P_{J-1}\lambda)\|_{\tau,\partial K},$$

so that we can use (5.33) to get

$$a_k((I - I_k P_{k-1})v, v) \le \|\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}\, P_{J-1}\lambda\|_c^2 + C\tau_K^{\max} h_K \, \|\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}(P_{J-1}\lambda)\|_K^2$$
$$\le C\|\vec{\mathfrak{Q}}^{\rm DG}\lambda + \vec{\nabla}\, P_{J-1}\lambda\|_c^2$$

and the theorem is proved for $\alpha = 1$ by using Lemma 5.12 with $s = 1$.

5.2 Numerical Experiments

In this section we report the performance of the multigrid solver for the HDG method, as we did previously for the HRT method in Chapter 3. We use the same mesh and problem settings as in § 3.3. Discretization errors in the finite element analysis are previously reported in § 4.3, so in this section we will only report observations on the multigrid convergence.

In order to check that the iteration error in the multigrid cycle converges at a rate independent of the mesh size h, we design the first experiment in a such way that we know the exact solution, as shown below:

1. We set $b = 0$, i.e., try to solve the equation system $Ax = 0$ with the exact solution $x = 0$.
2. The initial guess x_0 in the MG iteration on each multi-level space $M_k, k = 0, 1, \ldots, J$, is set to be $I_J I_{J-1} \cdots I_1 v$, where v is the linear combination of the global basis functions in M_0 with the coefficients $(1, \ldots, 1)$ (i.e., v takes the value 1 for all interior nodes, is linear on all mesh elements, is continuous across elements, and decreases to 0 on the boundary).
3. We use Gauss-Seidel iteration as the smoother. The number of smoothings, m_k, equals 2^{J-k} for HDG$(d = 0)$ and $4 \cdot 2^{J-k}$ for HDG$(d = 1)$, on m_k.
4. We stop iterations either when $\|x_i - x_{i-1}\|_a \le 10^{-8}\|x_0 - x\|_a$, or when the iteration count reaches the limit 99.

The experimental results presented in Table 5-1 and 5-2 for various values of τ illustrates the uniform convergence of the multigrid iteration we proved in Theorem 5.2. As we mentioned in Remark 5.3, the experiment results also indicate that the multigrid V-cycle converges uniformly even in the case of $d = 0$. Perhaps surprisingly the case $\tau = 1/h$ also yielded good multigrid convergence (unlike the case of the discretization error convergence – see § 4.3).

Table 5-1. The number of multigrid iterations ($m_k = 2^{J-k}$) for HDG(d=0) with different τ. * indicates divergence or an iteration count exceeding the limit.

mesh	0	1	2	3	4	5
$\tau = 1/2$	17	18	22	24	24	23
$\tau = 1$	17	18	22	24	24	23
$\tau = 2$	16	17	21	24	24	23
$\tau = 4$	16	16	21	24	24	23
$\tau = 8$	14	15	19	23	23	23
$\tau = 16$	19	22	18	21	22	23
$\tau = 32$	*	*	35	19	21	22
$\tau = 64$	*	*	*	50	19	20
$\tau = 128$	*	*	*	*	92	19
$\tau = 256$	*	*	*	*	*	*
$\tau = \|e\|$	17	18	22	25	24	24
$\tau = 1/\|e\|$	16	15	17	19	18	18

Table 5-2. The number of multigrid iterations ($m_k = 4 \cdot 2^{J-k}$) for HDG(d=1) with different τ.

mesh	0	1	2	3	4	5
$\tau = 1/2$	25	22	22	22	21	20
$\tau = 1$	24	22	22	21	21	20
$\tau = 2$	24	22	22	21	21	20
$\tau = 4$	23	21	22	21	21	20
$\tau = 8$	22	21	21	21	21	20
$\tau = 16$	20	20	20	20	20	20
$\tau = 32$	19	19	19	20	20	20
$\tau = 64$	18	18	18	18	19	19
$\tau = 128$	16	17	17	17	18	18
$\tau = 256$	16	16	16	16	17	17
$\tau = \|e\|$	25	22	22	22	21	20
$\tau = 1/\|e\|$	24	21	20	20	19	18

CHAPTER 6
CONCLUSION

This main product of this study is a fast and efficient iterative solution technique for hybridized finite element methods. We adopted the multigrid framework for designing such solvers. But at the initial stages, we were faced with the difficulties arising from the fact that the multilevel spaces in all hybridized methods were non-nested. In fact, the multilevel functions did not even share the same domain. We developed a multilevel framework in which this difficulty can be overcome. Within this framework, we analyzed two specific methods for which we are able to prove precise results. However, beyond the regime of proofs, the framework gives one the key ingredients to formulate a multigrid algorithm for any hybridized method. With the insights from our numerical experiments, we put forth these techniques with good confidence, even for the other hybridized methods.

Let us summarize the main theoretical results:

1. Considering the well known HRT mixed method first, we proved the uniform convergence of a multigrid V-cycle algorithm. Although multigrid for the mixed method has been studied previously by many authors, our algorithm is applied to the condensed hybrid bilinear form and is thus the most useful practically. A properly designed intergrid transfer operator was the key ingredient that opened the avenue towards a successful algorithm. We also provided negative results that certain seemingly plausible intergrid operators do not work.
2. Next, we studied the relatively recent class of DG methods called HDG methods. We proved new error estimates. We showed that the condition number of the HDG matrix grows like $O(h^{-2})$ as mesh size h tends to zero. The techniques developed here are anticipated to have utility in other DG methods of the hybrid type.
3. Finally, we showed that a multigrid algorithm gave an efficient solver for the matrix systems arising from the HDG method. We rigorously proved a convergence rate

independent of problem size, as well as provided numerical evidence of the efficacy of the algorithm. Typical multigrid analysis for DG methods usually only yields preconditioners, but with our carefully chosen intergrid transfer operator, we are able to prove the stronger result that the error reducing multigrid operator is a uniform contraction.

We have set apart for future research a number of questions that arose during this study. For example, although we have full proofs on the multigrid convergence for the HDG method with $d > 0$, the $d = 0$ part needs to be investigated further. Numerical experiments have strongly suggested that our algorithm also work for that case. The HDG $d = 0$ case is more than of academic interest as it uses spaces smaller than even the lowest order mixed method.

Extensions to other methods like HIP are also of interest. We have already proved a condition number estimate for the HIP method (Theorem 4.13). However, error estimates and multigrid convergence bounds remain unknown.

Finally, the modification of adaptive finite element techniques to hybridized methods is an interesting question of some significance. As HDG methods are gaining popularity, several groups are trying them out for practical problems. Treatment of practical problems require not only efficient solvers, but also adaptive refinements and a posteriori error estimators to automatically indicate which regions to refine. These are issues that currently remain open in the arena of HDG methods.

REFERENCES

[1] F. S. Acton, *Numerical Methods That Work*, Washington, DC: Math. Assoc. Amer.

[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] R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp., 36 (1981), pp. 35–51.

[4] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Springer-Verlag, New York, 1976.

[5] 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.

[6] J. H. Bramble, *Multigrid Methods*, no. 294 in Pitman research notes in mathematics series, Longman Scientific & Technical, Harlow, UK, 1993.

[7] J. H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp., 58 (1992), pp. 467–488.

[8] 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.

[9] 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.

[10] 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.

[11] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, no. 15 in Texts in Applied Mathematics, Springer-Verlag, New York, 1994.

[12] F. Brezzi, J. Douglas, and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235.

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

[14] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Company, Amsterdam, 1978.

[15] B. Cockburn, B. Dong, and J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp., 77 (2008), pp. 1887–1916.

[16] B. Cockburn, O. Dubois, J. Gopalakrishnan, and S. Tan, *Multigrid for the HDG method*, In Preparation, (2009).

[17] B. Cockburn and J. Gopalakrishnan, *A characterization of hybridized mixed methods for the Dirichlet problem*, SIAM J. Numer. Anal., 42 (2004), pp. 283–301.

[18] ——, *Error analysis of variable degree mixed methods for elliptic problems via hybridization*, Math. Comp., 74 (2005), pp. 1653–1677 (electronic).

[19] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM Journal on Numerical Analysis, 47 (2009), pp. 1319–1365.

[20] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Submitted, (2008).

[21] R. Courant, *Variational methods for the solution of problems of equilibrium and vibrations*, Bull. Amer. Math. Soc., 49 (1943), pp. 1–23.

[22] M. Dauge, *Elliptic boundary value problems on corner domains: smoothness and asymptotics of solutions*, no. 1341 in Lecture notes in mathematics, Springer-Verlag, Berlin, 1988.

[23] R. Falk and J. Osborn, *Error estimates for mixed methods*, RAIRO, 14 (1980), pp. 249–277.

[24] J. Gopalakrishnan, *A Schwarz preconditioner for a hybridized mixed method*, Computational Methods in Applied Mathematics, 3 (2003), pp. 116–134.

[25] J. Gopalakrishnan and S. Tan, *A convergent multigrid cycle for the hybridized mixed method*, Numer. Linear Algebra Appl., (2009), p. doi: 10.1002/nla.636.

[26] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, no. 24 in Monographs and Studies in Mathematics, Pitman Advanced Publishing Program, Marshfield, Massachusetts, 1985.

[27] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, vol. 95 of Applied Mathematical Sciences, Springer-Verlag, 1994.

[28] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436.

[29] R. B. Kellog, *Interpolation between subspaces of a Hilbert space*, Technical Note BN-719, University of Maryland, College Park, Maryland, November 1971.

[30] J. Nitsche, *Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens*, Numer. Math., 11 (1968), pp. 346–348.

[31] 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.

[32] J. N. Reddy, *An introduction to the finite element method*, McGraw-Hill, New York, 1993.

[33] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[34] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[35] 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.

BIOGRAPHICAL SKETCH

Shuguang Tan was born in China. He received his B.S. in Information Science from the Peking University in 2001. He came to U.S.A. in 2003 for continuing graduate study, and received his M.S in Statistics and Ph.D. in Mathematics concurrently from the University of Florida in the summer of 2009.