PRECONDITIONING THE MASS MATRIX FOR HIGH ORDER FINITE ELEMENT APPROXIMATION ON TETRAHEDRA*

MARK AINSWORTH† AND SHUAI JIANG†

Abstract. A preconditioner for the mass matrix arising from high order finite element discretisation on tetrahedra is presented and shown to give a condition number that is independent of both the mesh size and the polynomial order of the elements. The preconditioner is described in terms of a new, high-order basis which has the usual property whereby individual functions are associated with distinct geometric entities of the tetrahedron. It is shown that the basis enjoys the property that the resulting *mass matrix is spectrally equivalent to its own diagonal* with constants independent of h and p. Although the exposition is based on an explicit basis, the preconditioner can be applied to *any* choice of basis. In particular, the basis can be used to specify a basis independent Additive Schwarz Method (ASM), meaning that, in order to apply the preconditioner to an alternative basis, one only need implement an appropriate change of basis.

Key words. preconditioning mass matrix, polynomial extension theorem, high order finite element

AMS subject classifications. 65N30, 65N55, 65F08

1. Introduction.

In the p-version of the finite element method (p-FEM), one can obtain exponential rates of convergence [9, 31, 33], but the mass and stiffness matrices are generally poorly conditioned. The mass matrix for standard hierarchical bases have condition numbers that can grow as $\mathcal{O}(p^{12})$ [2, 16, 21, 24] while other bases such as Bernstein or Peano can exhibit even worse growth [20]. Large condition numbers can cause round off errors or mean that the cost of solving the linear systems unreasonably dominates, each of which potentially neutralizes the advantages of high order methods.

Effective preconditioners for the 3D stiffness matrix have been developed using domain decomposition [8, 36] methods. Depending on the sophistication and cost of the algorithm, condition numbers of the preconditioned stiffness matrix range from uniform to logarithmic growth in p [15, 22, 26, 30]. In contrast, until recently, there has been a dearth of preconditioners for the mass matrix on simplicial elements, with the exception of [4] which addressed the triangle case. In the present work, we develop a non-overlapping domain decomposition preconditioner for the mass matrix on tetrahedra which gives condition numbers *independent* of h and p. The preconditioner means that, e.g. in explicit time-stepping, one can increase p without fretting over the convergence of conjugate gradient.

Preconditioners for the mass matrix $\mathbf{M}$ for high-order C^0-conforming finite element methods have applications beyond just explicit and implicit time-stepping schemes. For instance, in the class of stationary equations, the singularly perturbed problem [6, 14], which arises in plate, beam and shell theories, gives rise to linear systems of the form $\mathbf{M} + \varepsilon^2\mathbf{S}$ where $\mathbf{S}$ is the stiffness matrix and $0 < \varepsilon \ll 1$. Similarly to the 2D case [5], our mass matrix preconditioner can be applied to the singularly perturbed system to give a condition number *independent of the parameter* ε on the optimal, single layer, anisotropic hp meshes which are advocated in [32] and shown

*Submitted to the editors DATE.

Funding: This work was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

†Department of Applied Mathematics, Brown University, Providence, RI. (mark_ainsworth@brown.edu, shuai_jiang@brown.edu)

to give robust exponential convergence in ε. By way of contrast, existing preconditioners [34] for anisotropic elements rely on a geometrically-graded mesh or tensor product elements in order to be robust in ε.

The preconditioner is described in terms of a new, high-order basis which has the usual property whereby individual functions are associated with distinct geometric entities of the tetrahedron. However, our basis enjoys the property that the resulting *mass matrix is spectrally equivalent to its own diagonal* with constants independent of h and p. Although the exposition is based on an explicit basis, the preconditioner can be applied to *any* choice of basis. In particular, the basis can be used to specify a basis-independent Additive Schwarz Method (ASM), meaning that, in order to apply the preconditioner to an alternative basis, one only needs to implement an appropriate change-of-basis.

In principle, the construction of an Additive Schwarz preconditioner for the mass matrix on tetrahedra should mirror the case for triangles [4]. In practice, however, one encounters a slew of difficulties associated with the stable decomposition of the face spaces which are not present in the the 2D case. In fact, even the choice of edge spaces and inner products turns out to be different from the 2D case owing to the need to decide how to extend the definition of the edge functions onto adjacent faces: in 2D one can rely on static condensation, but in 3D one is working with discrete trace norms defined implicitly by the Schur complement with respect to the interior functions in 3D. The net result is that the tetrahedral preconditioner is quite different from the case of the triangle. That said, our preconditioner for tetrahedra can be specialized to triangles to obtain a different preconditioner than the one developed in [4] which is *simpler* than the preconditioner in [4] and, in addition, gives a condition number roughly half the size.

The remainder of the paper is organized as follows. In section 2, we define the basis functions and state the main result. In section 3, we present illustrative numerical examples such as singularly perturbed problem and time-stepping. Finally in section 4, we prove the inequalities and polynomial extension lemmas needed for the main result.

2. Basis Definition and Main Result. Let T be the reference tetrahedron in $\mathbb{R}^3$ with vertices $v_1 = (-1,-1,-1), v_2 = (1,-1,-1), v_3 = (-1,1,-1),\ v_4 = (-1,-1,1)$, and let F_1 and E_1 be the face and edge given by

$$F_1 := T \cap \{z = -1\},$$
$$E_1 := T \cap \{z = -1\} \cap \{y = -1\}.$$

Let $p \geq 1$ be a given integer, and let $\mathbb{P}_p(D)$ be the space of polynomials of total degree p on a domain D. Let $X := \mathbb{P}_p(T)$, and $\lambda_i \in \mathbb{P}_1(T)$ for $i = 1,2,3,4$ be the barycentric coordinates of T associated with vertex v_i; i.e. $\lambda_i(v_j) = \delta_{ij}$.

We begin by introducing a particular basis for $\mathbb{P}_p(T)$ which, as usual, consists of functions associated with vertices, edges, faces and the interior of the tetrahedron. However, the actual choice of functions differs from those typically used in the literature.

2.1. Basis functions. The classical Jacobi polynomials [1] on $[-1,1]$ are denoted by $P_n^{(\alpha,\beta)}$, where n is the order of the polynomial and $\alpha, \beta > -1$ are weights, and satisfy

$$\int_{-1}^{1} \left(\frac{1-x}{2}\right)^{\alpha} \left(\frac{1+x}{2}\right)^{\beta} P_n^{(\alpha,\beta)}(x)^2\, dx = \frac{2(\alpha+n)!(\beta+n)!}{n!(\alpha+\beta+2n+1)(\alpha+\beta+n)!}.$$

For non-negative integers m, q, let $\Phi_q^{(m)}(x) \in \mathbb{P}_q([-1,1])$ be defined by

$$\Phi_q^{(m)}(x) := \frac{(-1)^q}{q+1} P_q^{(m,1)}(x), \tag{2.1}$$

and $\Xi_q \in \mathbb{P}_q([0,1]^2)$ be given by

$$\Xi_q(l_1, l_2) := P_q^{(2,2)}\left(\frac{2l_2}{l_1+l_2} - 1\right)(l_1+l_2)^q. \tag{2.2}$$

Interior Basis Functions. For $p \geq 4$, let

$$\omega_{ijk} := \lambda_1\lambda_2\lambda_3\lambda_4 \Xi_i(\lambda_1, \lambda_2) P_j^{(2i+5,2)}\left(\frac{2\lambda_3}{1-\lambda_4} - 1\right)(1-\lambda_4)^j \, P_k^{(2i+2j+8,2)}(2\lambda_4 - 1)$$

for $0 \leq i, j, k, i+j+k \leq p-4$. Note that ω_{ijk} vanishes on the boundary of T due to the factor $\lambda_1\lambda_2\lambda_3\lambda_4$. The set $\{\omega_{ijk}\}$ is an orthogonal basis for $X_I := X \cap H_0^1(T)$ with respect to the $L^2(T)$ inner product (see Lemma 4.1).

Face Basis Functions. For $p \geq 3$, the basis functions associated with the face F_1 are given by

$$\psi_{ij}^{(1)} := \lambda_1\lambda_2\lambda_3 \Xi_i(\lambda_1, \lambda_2) P_j^{(2i+5,2)}\left(\frac{2\lambda_3}{1-\lambda_4} - 1\right)(1-\lambda_4)^j \, \Phi_{p-3-i-j}^{(2i+2j+8)}(2\lambda_4 - 1)$$

for $0 \leq i, j, i+j \leq p-3$. In particular, the presence of the factor $\lambda_1\lambda_2\lambda_3$ means that these functions vanish on the remaining three faces. The basis functions on the other three faces F_k are defined in an analogous fashion to give the face spaces $X_{F_k} := \operatorname{span}\{\psi_{ij}^{(k)}\}$. The functions provide an orthogonal basis for X_{F_k} (e.g. $(\psi_{ij}^{(k)}, \psi_{mn}^{(k)}) \propto \delta_{ij,mn}$ where $(\cdot,\cdot)$ is the L^2 inner-product over T); see Lemma 4.1.

Edge Basis Functions. For $p \geq 2$, the basis functions associated with the edge E_1 are chosen as follows:

$$\chi_i^{(1)} := \lambda_1\lambda_2 \Xi_i(\lambda_1, \lambda_2) \frac{q_i(\lambda_3, \lambda_4) + q_i(\lambda_4, \lambda_3)}{2}, \qquad 0 \leq i \leq p-2,$$

where the function q_i is given by

$$q_i(l_1, l_2) := \Phi_j^{(2i+5)}\left(\frac{2l_1}{1-l_2} - 1\right)(1-l_2)^j \, \Phi_{p-2-i-j}^{(2i+2j+6)}(2l_2 - 1) \tag{2.3}$$

with $j = \lfloor (p-i-2)/2 \rfloor$. The basis functions on the remaining edges E_k are defined analogously to give the edge spaces $X_{E_k} := \operatorname{span}\{\chi_i^{(k)}\}$.

The edge basis functions have the following properties:

1. locally supported: vanish on the two faces which do not contain edge E_1 (owing to the factor $\lambda_1\lambda_2$);
2. symmetry: the values on the two non-zero faces satisfy the condition that $\chi(r,s,t,0) = \chi(r,s,0,t)$ for all r, s, t;
3. orthogonality: $(\chi_i^{(k)}, \chi_j^{(k)}) \propto \delta_{ij}$ (see Lemma 4.1).

Vertex Basis Functions. The function associated with the vertex v_1 is given by

$$\varphi_1 := \frac{1}{3}\lambda_1 \left(q(\lambda_2, \lambda_3, \lambda_4) + q(\lambda_3, \lambda_4, \lambda_2) + q(\lambda_4, \lambda_2, \lambda_3)\right)$$

where

$$q(l_1, l_2, l_3) := \Phi_i^{(2)}\left(\frac{2l_1}{1-l_2-l_3} - 1\right)(1-l_2-l_3)^i\,\Phi_j^{(2i+3)}\left(\frac{2l_2}{1-l_3} - 1\right) \times (1-l_3)^j\,\Phi_{p-1-i-j}^{(2i+2j+4)}(2l_3 - 1), \tag{2.4}$$

with $i = \lfloor \frac{p}{2} \rfloor$ and $j = \lfloor \frac{i}{2} \rfloor$. The basis functions on the remaining vertices are defined in an analogous manner to give the vertex spaces $X_{V_k} := \text{span}\{\varphi_k\}$.

The vertex basis functions have the following properties:

1. local support: $\varphi_1(v_1) = 1$ and vanishes at the remaining vertices;
2. symmetry: the values on the three non-zero faces satisfy the condition that $\varphi_1(r, s, 0, 0) = \varphi_1(r, 0, s, 0) = \varphi_1(r, 0, 0, s)$ for all r, s.

It is not difficult to see that the basis functions are linearly independent and a simple counting argument shows that the union of the sets gives a basis for X.

Basis Functions on a Mesh. Let Ω be a bounded three-dimensional domain, and let $\mathcal{P}$ be a partitioning of Ω into the union of disjoint tetrahedra such that the intersection of any two distinct elements is either a single common vertex, edge or face. Each element $K \in \mathcal{P}$ is the image of the reference element T under a (possibly non-affine) map $\mathcal{F}_K$ such that there exists positive constants θ, Θ such that the Jacobian $D\mathcal{F}_K$ satisfies

$$\theta|K| \leq |D\mathcal{F}_K(x)| \leq \Theta|K| \qquad \forall x \in K. \tag{2.5}$$

It is worth noting that this condition does not place constraints on the shape regularity of the mesh, and, in particular, allows for "needle" or "slab" elements.

The basis functions on an element $K \in \mathcal{P}$ are defined to be pull-backs using the map $\mathcal{F}_K$ in the usual manner, e.g.

$$\varphi_{1,K}(x) := \varphi_1(\mathcal{F}_K^{-1}(x)), \qquad x \in K.$$

The fact that the basis functions are associated with vertices, edges and faces, together with the symmetry properties means that enforcing global conformity follows the same procedure for hierarchic bases. In particular, one needs to number the degrees of freedom in a systematic manner to ensure that the edge and face basis functions will be oriented correctly. The standard finite element sub-assembly gives the global mass matrix

$$\mathbf{M} = \sum_{K \in \mathcal{P}} \mathbf{\Lambda}_K \mathbf{M}_K \mathbf{\Lambda}_K^T$$

where $\mathbf{\Lambda}_K$ is the local assembly matrix and $\mathbf{M}_K$ is the element mass matrix expressed using the above basis. For more details about the assembly process, see [3].

2.2. Main result. The main result states that the diagonal of the mass matrix is spectrally equivalent to the full matrix:

THEOREM 2.1. *Suppose that the basis is chosen as in subsection 2.1. Then, there exists constants* τ, Υ *independent of* h, p *such that*

$$\tau \operatorname{diag}(\mathbf{M}) \leq \mathbf{M} \leq \Upsilon \operatorname{diag}(\mathbf{M}).$$

Proof. Let $\hat{\mathbf{M}}$ be the mass matrix on the reference element T, then (2.5) implies that

$$\theta |K| \hat{\mathbf{M}} \leq \mathbf{M}_K \leq \Theta |K| \hat{\mathbf{M}}. \tag{2.6}$$

We shall show below that the following condition holds with constants c, C independent of p:

$$c \operatorname{diag}(\hat{\mathbf{M}}) \leq \hat{\mathbf{M}} \leq C \operatorname{diag}(\hat{\mathbf{M}}). \tag{2.7}$$

Then, sub-assembly together with (2.6) and (2.7) shows that

$$\begin{aligned} c \operatorname{diag}(\mathbf{M}) &= c \sum_{K \in \mathcal{P}} \mathbf{\Lambda}_K \operatorname{diag}(\mathbf{M}_K) \mathbf{\Lambda}_K^T \leq c \sum_{K \in \mathcal{P}} |K| \mathbf{\Lambda}_K \operatorname{diag}\left(\hat{\mathbf{M}}\right) \mathbf{\Lambda}_K^T \\ &\leq \sum_{K \in \mathcal{P}} |K| \mathbf{\Lambda}_K \hat{\mathbf{M}} \mathbf{\Lambda}_K^T \leq C \sum_{K \in \mathcal{P}} |K| \mathbf{\Lambda}_K \operatorname{diag}\left(\hat{\mathbf{M}}\right) \mathbf{\Lambda}_K^T \\ &\leq C \sum_{K \in \mathcal{P}} \mathbf{\Lambda}_K \operatorname{diag}(\mathbf{M}_K) \mathbf{\Lambda}_K^T = C \operatorname{diag}(\mathbf{M}) \end{aligned} \tag{2.8}$$

where we dropped the dependence on θ, Θ.

It remains to show that condition (2.7) holds: that is, there exists constants c, C independent of p such that

$$c \vec{u}^T \operatorname{diag}(\hat{\mathbf{M}}) \vec{u} \leq \vec{u}^T \hat{\mathbf{M}} \vec{u} \leq C \vec{u}^T \operatorname{diag}(\hat{\mathbf{M}}) \vec{u}, \qquad \forall \vec{u}.$$

The result is trivial for $p = 1, 2$ and 3 by equivalence of norms on the spaces $\mathbb{P}_1, \mathbb{P}_2$ and $\mathbb{P}_3$. It suffices to consider the case $p \geq 4$.

Let $u \in X$ be the function corresponding to $\vec{u}$ so that $\vec{u}^T \hat{\mathbf{M}} \vec{u} = \|u\|^2$ where $\|\cdot\|$ is the standard L^2 norm over T. The vector $\vec{u}$ can be decomposed as follows:

$$\vec{u} = [\vec{u}_I, \vec{u}_{F_1}, \ldots, \vec{u}_{F_4}, \vec{u}_{E_1}, \ldots, \vec{u}_{E_6}, \vec{u}_{V_1}, \ldots, \vec{u}_{V_4}]$$

where $\vec{u}_I$ corresponds to the coefficients of the interior basis functions ω_{ijk} or, equally well, a function $u_I \in X_I$ etc. This partitioning induces a partitioning of the mass matrix into subblocks. Moreover, the orthogonality of the basis functions *within* each block (but not necessarily *between* different blocks) means that

$$\operatorname{diag}(\hat{\mathbf{M}}) = \begin{bmatrix} \hat{\mathbf{M}}_I & & & \\ & \hat{\mathbf{M}}_{F_1} & & \\ & & \ddots & \\ & & & \hat{\mathbf{M}}_{V_4} \end{bmatrix}.$$

Thus,

$$\vec{u}^T \operatorname{diag}(\hat{\mathbf{M}}) \vec{u} = \|u_I\|^2 + \sum_{i=1}^{4} \|u_{F_i}\|^2 + \sum_{i=1}^{6} \|u_{E_i}\|^2 + \sum_{i=1}^{4} \|u_{V_i}\|^2$$

where $u_I \in X_I$, $u_{F_i} \in X_{F_i}$, $u_{E_i} \in X_{E_i}$ and $u_{V_i} \in X_{V_i}$.

Condition (2.7) hence reduces to showing that for all $u \in X$, there exist positive constants c, C independent of p such that

$$c\left(\|u_I\|^2 + \sum_{i=1}^{4}\|u_{F_i}\|^2 + \sum_{i=1}^{6}\|u_{E_i}\|^2 + \sum_{i=1}^{4}\|u_{V_i}\|^2\right) \le \|u\|^2 \le C\left(\|u_I\|^2 + \sum_{i=1}^{4}\|u_{F_i}\|^2 + \sum_{i=1}^{6}\|u_{E_i}\|^2 + \sum_{i=1}^{4}\|u_{V_i}\|^2\right). \tag{2.9}$$

The upper-bound follows at once thanks to the triangle inequality. The proof of the lower bounds is less straight forward and relies on a number of technical estimates whose proofs are postponed to section 4.

Lemma 4.4 and the fact that $\|u\|_\infty \le Cp^3\|u\|$ [38] gives the following bound on the vertex components:

$$\|u_{V_i}\| = \|u(v_i)\varphi_i\| \le \|\varphi_i\|\|u\|_\infty \le C\|u\|, \qquad i = 1, \dots, 4.$$

Now, by Lemma 4.5, we obtain

$$\|u_{E_i}\| \le C\left\|u - \sum_{i=1}^{4} u_{V_i}\right\| \le C\|u\|, \qquad i = 1, \dots 6.$$

We next apply Corollary 4.7 to each individual face to obtain

$$\|u_{F_i}\| \le C\left\|u - \sum_{i=1}^{4} u_{V_i} - \sum_{i=1}^{6} u_{E_i}\right\| \le C\|u\|, \qquad i = 1, 2, 3, 4.$$

Finally, a bound for u_I is an easy consequence of the triangle inequality

$$\|u_I\| \le C\left\|u - \sum_{i=1}^{4} u_{V_i} - \sum_{i=1}^{6} u_{E_i} - \sum_{i=1}^{4} u_{F_i}\right\| \le C\|u\|. \qquad \square$$

Collecting these estimates establishes the lower bound in (2.9).

3. Numerical Examples.

3.1. Preconditioned mass matrix. We first illustrate Theorem 2.1 for a single element. The left side of Figure 1 shows the condition number of the preconditioned mass matrix on the reference tetrahedron. As predicted by Theorem 2.1, the condition numbers remain bounded as p is increased.

To illustrate the h independence of the preconditioned system, we consider the two meshes illustrated in Figure 2. The right side of Figure 1 shows the condition number of

$$\mathbf{M}_s := \mathbf{P}^{-1/2}\mathbf{M}\mathbf{P}^{-1/2}$$

where $\mathbf{M}$ is the global mass matrix on the cube and $\mathbf{P} = \text{diag}(\mathbf{M})$ on these meshes. It is observed that the condition numbers on the refined meshes track the condition numbers obtained on a single tetrahedron as suggested by (2.8).

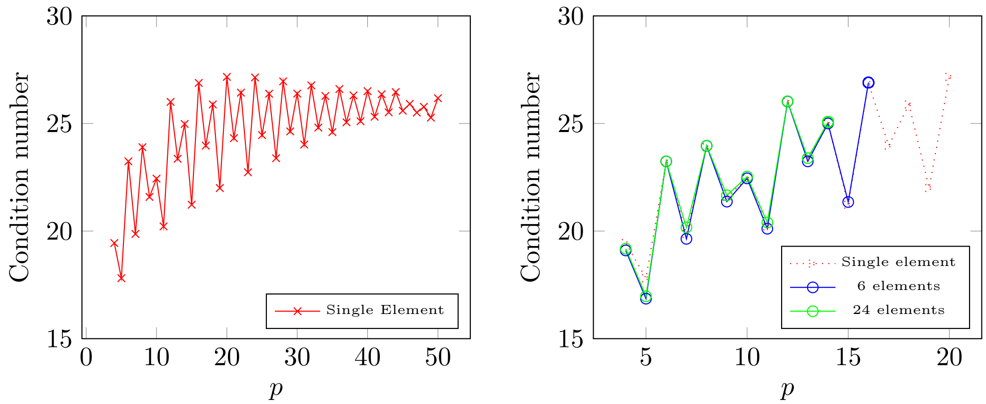

FIG. 1. *Figure illustrates the condition number of the preconditioned mass matrix on a meshes of six elements, 24 elements and on a mesh consisting of a single element. The bounded condition number of the preconditioned system is in agreement with Theorem 2.1.*

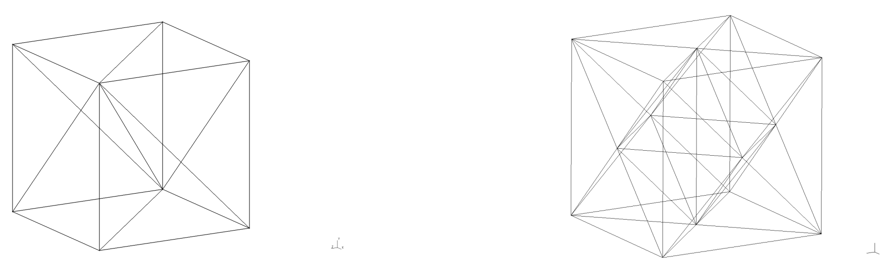

FIG. 2. *Figure illustrating the two meshes on the cube. The mesh on the left contains six elements and the mesh on the right contains 24 elements.*

3.2. Singularly Perturbed Problem. The utility of the preconditioner is not confined to the pure mass matrix. Consider the following problem

$$\begin{aligned} u - \varepsilon^2 \Delta u &= f, \qquad x \in \Omega, \\ u &= 0, \qquad x \in \partial\Omega, \end{aligned} \tag{3.1}$$

where $0 < \varepsilon \ll 1$ and $f \in L^2(\Omega)$ which is prototypical of several class of problem arising in mechanics [6,14]. The p-version Galerkin discretization of (3.1) leads to an algebraic problem of the form

$$(\mathbf{M} + \varepsilon^2 \mathbf{S})\vec{u} = \vec{f} \tag{3.2}$$

where $\mathbf{S}$ is the stiffness matrix and $\vec{f}$ is the load vector corresponding to f.

Solutions to (3.1) generally exhibit boundary layers which become sharper as $\varepsilon \to 0$; see Figure 3 for a plot of the solution for $f = 1$. If the order of the finite element method p is large enough so that $\mathcal{O}(p\varepsilon) \geq 1$, then one obtains exponential convergence in p on a quasi-uniform mesh [23]. If $\varepsilon \ll 1$, then it is unrealistic to choose the degree $p = \mathcal{O}(\varepsilon^{-1}) \gg 1$. Instead, a single layer of anisotropic elements of width $\mathcal{O}(p\varepsilon)$ around the boundary suffices [23] to give robust exponential convergence in p *independent of* ε. Whilst this restores the accuracy of the resulting approximations, an undesirable side-effect of the anisotropic elements is that the condition number of

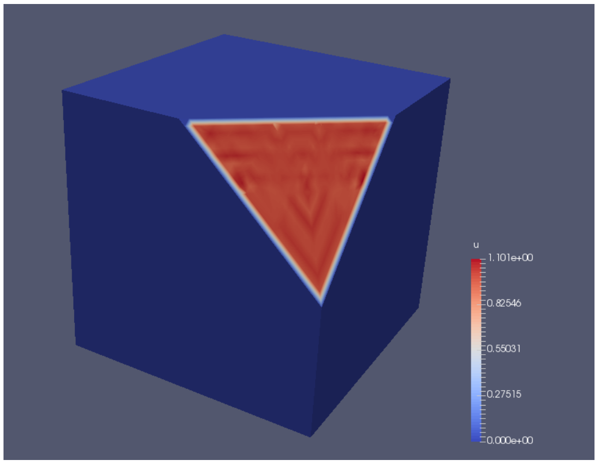

Fig. 3. *Cross-section of the solution to* (3.1) *for* $\varepsilon^2 = 10^{-4}$ *and* $p = 10$ *on a corner of the cube showing the presence of a boundary layer.*

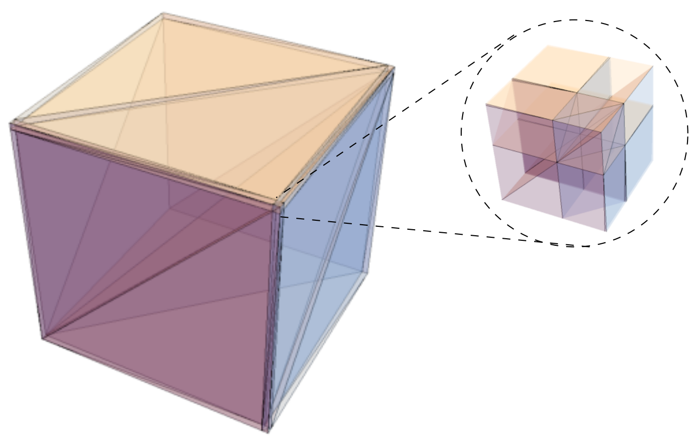

Fig. 4. *Figure illustrating the mesh used to approximate the singularly perturbed problem on an octant of the cube. The inset shows the submesh of elements in the corner. Note the needle and slab elements of width* $\mathcal{O}(p\varepsilon)$ *encompassing the boundary of the cube.*

(3.2) grows rapidly as $\varepsilon \to 0$. This means that the system (3.2) becomes increasingly difficult to solve unless a preconditioner is used. Toselli and Vasseur [34,35] developed a domain decomposition preconditioner for tensor product elements which results in a condition number independent of ε and growing as $1 + \log^2 p$. Unfortunately, the analysis of Toselli and Vasseur relies strongly on a tensor product structure and only holds on a geometrically graded mesh. In particular, it does not apply to the boundary layer mesh of [23] described above nor to meshes of tetrahedra. There are effectively *no* existing preconditioners which are robust in the aspect ratio ε on simplices. However, it turns out that using a mass matrix as a preconditioner gives a condition number independent of ε with a $\mathcal{O}(p^2)$ growth on the boundary layer mesh described above.

A similar idea was first explored in [5] in the two dimensional case. We shall need the following result:

LEMMA 3.1. *Let K be a slab or needle tetrahedron with the smallest side length of size $p\varepsilon \ll 1$, then for all polynomials $u \in \mathbb{P}_p(K)$, there exists a constant C independent of ε, p such that*

$$\|\nabla u\|_K^2 \leq C\frac{p^2}{\varepsilon^2}\|u\|_K^2 .$$

Proof. Consider the case of the slab tetrahedron first. Without loss of generality, let K be the slab tetrahedron defined by the vertices $(0,0,0)$, $(p\varepsilon,0,0)$, $(0,1,0)$, $(0,0,1)$ and let $\hat{K}$ be the tetrahedron with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$.

Given $u \in \mathbb{P}_p(K)$, let $\hat{u}(\hat{x},\hat{y},\hat{z}) = u(p\varepsilon\hat{x},\hat{y},\hat{z})$ be the polynomial defined on $\hat{K}$, then by a change of variables

$$\begin{aligned}
\|\nabla u\|_K^2 &= \int_K |\nabla u|^2 \, dxdydz \\
&= \int_{\hat{K}} \frac{1}{(p\varepsilon)^2}(\partial_{\hat{x}}\hat{u}(\hat{x},\hat{y},\hat{z}))^2 + (\partial_{\hat{y}}\hat{u}(\hat{x},\hat{y},\hat{z}))^2 + (\partial_{\hat{z}}\hat{u}(\hat{x},\hat{y},\hat{z}))^2 p\varepsilon \, d\hat{x}d\hat{y}d\hat{z} \\
&\leq \frac{1}{p\varepsilon}\int_{\hat{K}} |\hat{\nabla}\hat{u}|^2 \, d\hat{x}d\hat{y}d\hat{z} \\
&\leq \frac{C_S p^3}{\varepsilon}\int_{\hat{K}} \hat{u}^2 \, d\hat{x}d\hat{y}d\hat{z} \\
&= \frac{C_S p^3}{\varepsilon}\int_K u^2 \, \frac{1}{p\varepsilon} dxdydz = C_S\frac{p^2}{\varepsilon^2}\|u\|_K^2
\end{aligned}$$

where we used the standard Schmidt's inequality $\|\nabla u\|_{\hat{K}}^2 \leq C_S p^4 \|u\|_{\hat{K}}^2$ on the reference element $\hat{K}$ [10, 25].

The proof for the needle element follows similarly by using the transformation $\hat{u}(\hat{x},\hat{y},\hat{z}) = u(p\varepsilon\hat{x}, p\varepsilon\hat{y},\hat{z})$. ◻

The above lemma in conjunction with Theorem 2.1 gives rise to the following bound

$$c\,\mathrm{diag}(\mathbf{M}) \leq \mathbf{M} + \varepsilon^2\mathbf{S} \leq \left(1 + C\varepsilon^2\frac{p^2}{\varepsilon^2}\right)\mathbf{M} \leq Cp^2\mathrm{diag}(\mathbf{M}) \tag{3.3}$$

on a mesh where a layer of slab and needle elements of width $p\varepsilon$ are placed along the boundary; see Figure 4 for an figure of the mesh used on an octant of the cube. Equation (3.3) shows that using the mass matrix preconditioner to precondition the system (3.2) results in a condition number that grows as $\mathcal{O}(p^2)$ but, crucially, remains *independent* of ε, even on an unstructured mesh.

To illustrate the overall effectiveness of the approach of using the boundary layer mesh from [23] alongside the mass matrix preconditioner, we consider problem (3.1) with $f = 1$ and $\Omega = (-100, 100)^3$. Due to symmetry of the problem, it suffices to only consider the octant of the cube given by $(0, 100)^3$ which we illustrated in Figure 4. The condition number of the preconditioned matrices

$$\mathrm{diag}(\mathbf{M})^{-1/2}\left(\mathbf{M} + \varepsilon^2\mathbf{S}\right)\mathrm{diag}(\mathbf{M})^{-1/2}$$

is reported in Table 1 where it is seen that the condition number is indeed bounded independent of ε.

TABLE 1
Condition number of the singularly perturbed matrices obtained using the preconditioner for the pure mass matrix. Observe the condition number exhibits moderate growth in p but remains bounded independent of ε.

ε^2	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$
1e-1	16.99	19.76	28.56	33.88	59.27	83.03
1e-3	22.61	21.17	30.65	30.02	39.20	39.04
1e-5	23.24	22.09	32.75	31.41	42.32	40.15
1e-7	23.31	22.25	33.08	31.67	42.78	40.38
1e-9	23.31	22.27	33.11	31.70	42.83	40.41

3.3. Time-Stepping. Finally, we discuss the application of the preconditioner to time-stepping problems. Let

$$\mathbf{A}(\mu,\nu) := \mu\mathbf{M} + \nu\Delta t\mathbf{S}.$$

For a fully explicit scheme $\nu = 0$, and Theorem 2.1 implies that the preconditioner will be uniform in the polynomial order p. For a implicit scheme $\nu > 0$, we once again take advantage of Schmidt's inequality, which states that there exists a constant C_S independent of h, p such that $\mathbf{S} \le C_S\frac{p^4}{h^2}\mathbf{M}$, to deduce that

$$\mu\mathbf{M} \le \mathbf{A}(\mu,\nu) \le (\mu + C_S\frac{p^4}{h^2}\nu\Delta t)\mathbf{M} \le 2\max\left(\mu, C_S\frac{p^4}{h^2}\nu\Delta t\right)\mathbf{M}.$$

In other words, preconditioning using the diagonal of the mass matrix gives

$$\operatorname{cond}(\tilde{\mathbf{A}}(\mu,\nu)) \le \frac{2\Upsilon}{\tau}\max\left(1, C_S\frac{p^4\nu\Delta t}{h^2\mu}\right) \tag{3.4}$$

where $\tilde{\mathbf{A}}(\mu,\nu) = \operatorname{diag}(\mathbf{M})^{-1/2}\mathbf{A}(\mu,\nu)\operatorname{diag}(\mathbf{M})^{-1/2}$ and τ, Υ are the constants from Theorem 2.1; in practice one does not see the $\mathcal{O}(p^4)$ growth owing to the small value of the multiplicative factor $C_S\nu\Delta t/\mu$.

For a concrete example, consider a system of nonlinear reaction-diffusion equations [13] which exhibits pattern formation [27]:

$$\begin{aligned} \frac{\partial u}{\partial t} &= -uv^2 + \alpha(1-u) + d_u\Delta u \\ \frac{\partial v}{\partial t} &= uv^2 - (\alpha+\beta)v + d_v\Delta v \end{aligned} \qquad (x,y) \in \Omega, t > 0, \tag{3.5}$$

where $\alpha = .05, \beta = .02, d_u = 2\times 10^{-5}, d_v = 10^{-5}$ and Ω a hemisphere with radius 1. Figure 7 illustrates the solution u at $t = 1500$. It is commonplace in applications for the diffusion coefficients to be significantly smaller in magnitude than the reaction terms. For example, the Brusselator system arising in computational chemistry considered in [17, 37] or the Schnakenberg system arising in developmental biology considered in [28, 39] each have diffusion coefficients at least two orders of magnitude smaller than the corresponding reaction factors.

Using a standard Galerkin approximation in the spatial dimensions and an IMEX scheme [28] for the temporal dimension, one arrives at the follow linear systems:

$$\begin{aligned} \frac{\mathbf{M}\vec{u}^{n+1} - \mathbf{M}\vec{u}^{n}}{\Delta t} &= -\vec{g}^n + \alpha \vec{1} - \alpha \mathbf{M}\vec{u}^{n+1} - \frac{d_u}{2}\left(\mathbf{S}u^{n+1} + \mathbf{S}u^n\right) \\ \frac{\mathbf{M}\vec{v}^{n+1} - \mathbf{M}\vec{v}^{n}}{\Delta t} &= \vec{g}^n - (\alpha+\beta) \mathbf{M}\vec{v}^{n+1} - \frac{d_v}{2}\left(\mathbf{S}v^{n+1} + \mathbf{S}v^n\right) \end{aligned} \tag{3.6}$$

where $\vec{u}^n, \vec{v}^n$ is the finite element approximation at time step n and $\vec{g}^n$ is the nonlinear moment associated with uv^2 at time step n. An IMEX scheme is chosen since the diffusion operator is stiff and necessitates prohibitively small time steps were an explicit method to be chosen.

The first equation of (3.6) involves inverting the matrix $\mathbf{A}\left(1 + \alpha\Delta t, d_u/2\right)$ at each time step. Since $\mu \gg \nu$ and numerical evidence suggests that the constant $C_S < \frac{1}{5}$ [25], the constant in front of the $\mathcal{O}(p^4)$ growth in (3.4) is quite small. In Figure 5, we show the condition number of $\tilde{\mathbf{A}}\left(1 + \alpha\Delta t, d_u/2\right)$ with different Δt and order p. In practice, one generally chooses Δt depending on p, but for illustrative purposes here, we vary Δt and p independently. Note that the condition number for $p \leq 10$ does not yet attain the asymptotic $\mathcal{O}(p^4)$ growth even for artificially large values of Δt. Results for the case $\Delta t = 5$ also exhibit a transition from constant condition number to a slight growth with p as predicted by (3.4).

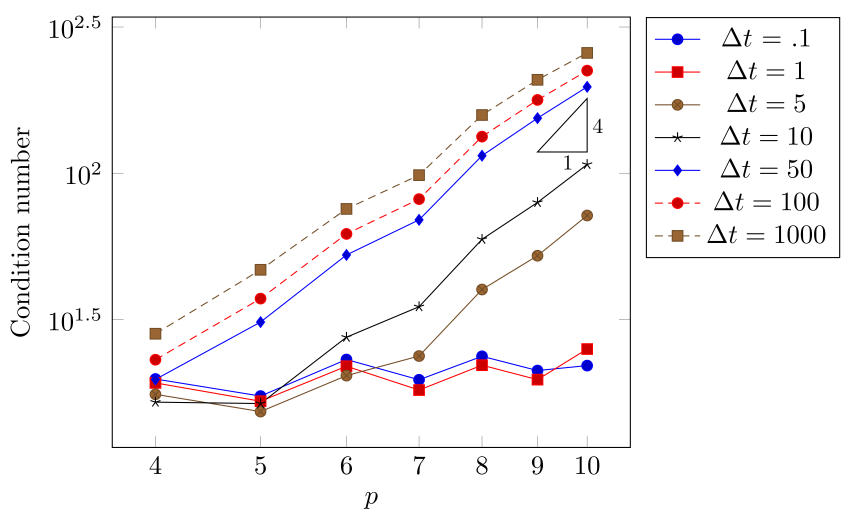

FIG. 5. *Figure illustrating the condition number of the preconditioned system arising from the discretization of the reaction-diffusion system on the hemisphere consisting of* 60 *elements. Note that we do not yet observe the* $\mathcal{O}(p^4)$ *growth for* $p \leq 10$ *even for very large* Δt.

The practical value of the preconditioner is illustrated in Table 2 where we display the [min, median, max] iteration count resulting from using preconditioned conjugate gradient (PCG) to perform time stepping for the Gray-Scott example to $t = 100$ with $\Delta t = 1$ for the v variable. The number of iterations is seen to remain bounded as suggested by the condition numbers depicted in Figure 5. Figure 6 shows the residuals of PCG at $t = 0$ for the v variable which are seen to decrease at a steady rate.

3.4. Application to the Nonsymmetric Systems. The mass matrix preconditioner is also useful in cases where the linear system is not symmetric. For instance,

TABLE 2

Table displays the [min, median, max] iteration count of PCG applied to the system $\tilde{\mathbf{A}}\left(1+\alpha\Delta t, d_u/2\right)$ *resulting from the IMEX scheme (3.6) for a period of 100 seconds with* $\Delta t = 1$ *on 60 elements for the reaction diffusion equation on the half-hemisphere.*

p	Preconditioned Iteration Count
4	[13, 14, 18]
6	[12, 13, 17]
8	[11, 11, 15]
10	[7, 10, 15]

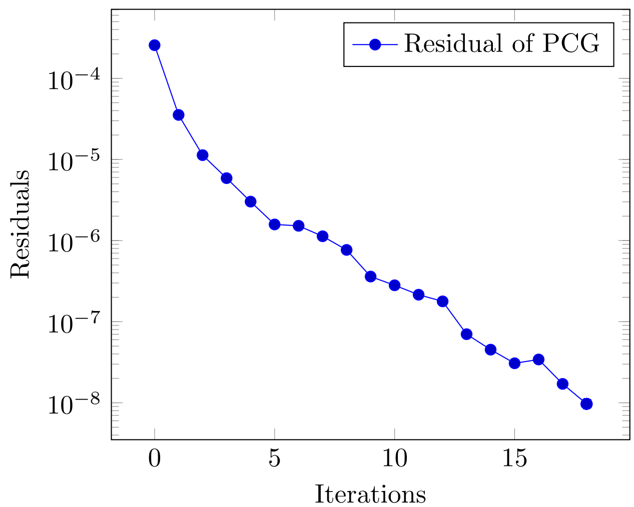

FIG. 6. *Plot of the residuals resulting from the preconditioned conjugate gradient method applied to the Gray-Scott example with* $p = 6$ *on the hemisphere at* $t = 0$ *for the* v *variable.*

consider the linear advection equation

$$\frac{\partial u}{\partial t} = \nu \cdot \nabla u, \qquad (x, y) \in \Omega, t > 0 \tag{3.7}$$

subject to $u = 0$ on $\partial\Omega, t > 0$ and $u(x, 0) = u_0(x)$ in Ω, where ν is a velocity field. For simplicity, we consider a standard Galerkin approximation in space and backward Euler in time. The resulting linear system is

$$\mathbf{B}\vec{u}^{n+1} = \mathbf{M}\vec{u}^{n}, \qquad \mathbf{B} := \mathbf{M} + \Delta t\mathbf{C} \tag{3.8}$$

where $\vec{u}^n$ is the finite element approximation at time n, $\mathbf{C}$ is the convective matrix with entries $\mathbf{C}_{ij} = (\varphi_i, \nu \cdot \nabla\varphi_j)$ and φ_i, φ_j are the basis functions. Observe that $\mathbf{M}$ is SPD whilst $\mathbf{C}$ is skew-symmetric and thus has a purely imaginary spectrum. Moreover, we have for any vector $\vec{u}$

(3.9)

$$|\vec{u}^T\mathbf{C}\vec{u}| \leq |(u, \nu \cdot \nabla u)| \leq \|\nu\|_{L^\infty}\|u\|\|\nabla u\| \leq \frac{C_S p^2}{h}\|\nu\|_{L^\infty}\|u\|^2 = \frac{C_S p^2}{h}\|\nu\|_{L^\infty}\,\vec{u}^T\mathbf{M}\vec{u}$$

where C_S is the constant arising from Schmidt's inequality. In particular, this means that if $\Delta t \ll C\frac{h}{p^2}$, then $\mathbf{B} \sim \mathbf{M}$ which suggests using $\mathbf{M}$ as a preconditioner for $\mathbf{B}$.

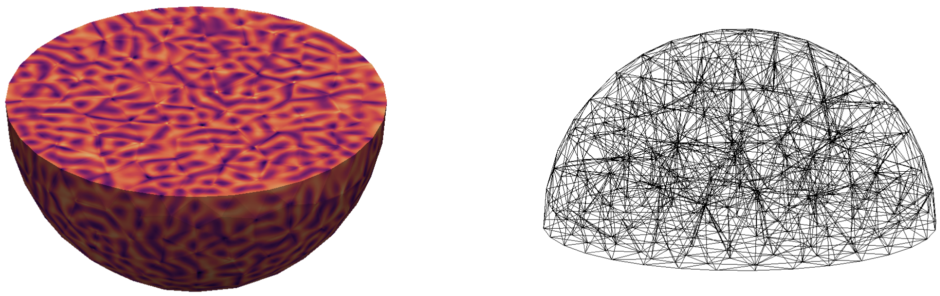

FIG. 7. *Plot of u from above in the Gray-Scott equations (3.5) with $p = 6$ (left) on a mesh of the hemisphere with 1159 elements (right) at $t = 1500$ with $\Delta t = 1$.*

The resulting preconditioned matrix

$$\hat{\mathbf{B}} := \mathbf{M}^{-1/2}\mathbf{B}\mathbf{M}^{-1/2} = \mathbf{I} + \Delta t\mathbf{M}^{-1/2}\mathbf{C}\mathbf{M}^{-1/2}$$

has eigenvalues which lie on the segment $S = [1 - i\Lambda, 1 + i\Lambda] \subset \mathbb{C}$ with $\Lambda = C\Delta t\frac{p^2}{h}$. If GMRES [29] is used to solve systems involving the matrix $\hat{\mathbf{B}}$, then, thanks to [12, Corollary 2.8] and [29, Proposition 6.32], the residual at the k-th iteration is bounded by

$$\|\vec{r}_k\| \leq \frac{\Lambda}{\sqrt{1+\Lambda^2}}\left(\frac{\Lambda}{1+\sqrt{1+\Lambda^2}}\right)^{k-1}\|\vec{r}_0\| \tag{3.10}$$

where $\vec{r}_0$ is the initial residual. This estimate shows that if Δt is small, e.g. such that $\Lambda \leq 1$, then the quantity $\frac{\Lambda}{1+\sqrt{1+\Lambda^2}} < 1/2$ and one obtains rapid convergence. In practice, one chooses $\Delta t \sim h/p$ so that $\Lambda \sim \mathcal{O}(p)$ meaning that GMRES will converge at a rate which degenerates slowly with the order p.

The above discussion suggests using the preconditioner for the mass matrix as a preconditioner for $\mathbf{B}$, giving rise to the preconditioned operator

$$\tilde{\mathbf{B}} := \mathrm{diag}(\mathbf{M})^{-1/2}\mathbf{B}\mathrm{diag}(\mathbf{M})^{-1/2} = \mathbf{M}_S + \Delta t\mathbf{C}_S \tag{3.11}$$

with $\mathbf{M}_S = \mathrm{diag}(\mathbf{M})^{-1/2}\mathbf{M}\mathrm{diag}(\mathbf{M})^{-1/2}$ and $\mathbf{C}_S = \mathrm{diag}(\mathbf{M})^{-1/2}\mathbf{C}\mathrm{diag}(\mathbf{M})^{-1/2}$. The estimate (3.9) along with Theorem 2.1 reveals that

$$|\vec{u}^T\mathbf{C}_S\vec{u}| \leq \frac{C_S p^2}{h}\|\nu\|_{L^\infty}\,\vec{u}^T\mathrm{diag}(\mathbf{M})^{-1/2}\mathbf{M}\mathrm{diag}(\mathbf{M})^{-1/2}\vec{u} \leq \frac{C\Upsilon p^2}{h}\vec{u}^T\vec{u}$$

where Υ is the upper bound arising in Theorem 2.1. Consequently, using the fact that $\rho(\mathbf{A}) = \|\mathbf{A}\|$ for $\mathbf{A}$ a normal matrix where $\rho(\cdot)$ is the spectral radius of a matrix, we have

$$\left\|\tilde{\mathbf{B}}\right\| \leq \|\mathbf{M}_S\| + \Delta t\|\mathbf{C}_S\| = \rho(\mathbf{M}_S) + \Delta t\rho(\mathbf{C}_S) \leq \Upsilon(1 + C\Delta t p^2/h)$$

and $\lambda_{\min}(\tilde{\mathbf{B}} + \tilde{\mathbf{B}}^T) \geq 2\tau$ where τ is the lower bound arising in Theorem 2.1. Finally, Elman [7,11] gives the following bound for the convergence of GMRES for the matrix $\tilde{\mathbf{B}}$,

$$\|\vec{r}_k\| \leq \sin^k(\beta)\|\vec{r}_0\|$$

TABLE 3
Iteration count of using GMRES to solve the preconditioned system $\tilde{\mathbf{B}}$ and unpreconditioned system $\mathbf{B}$. Using the preconditioner greatly reduces the iteration count in all cases.

p	$\Delta t = 0.001$		$\Delta t = 0.01$		$\Delta t = 0.1$	
	$\tilde{\mathbf{B}}$	$\mathbf{B}$	$\tilde{\mathbf{B}}$	$\mathbf{B}$	$\tilde{\mathbf{B}}$	$\mathbf{B}$
4	20	107	22	104	100	365
5	28	285	22	243	186	1438
6	25	855	37	611	213	6699
7	24	2380	41	1798	269	26573
8	31	4582	58	3060	286	99102
9	27	15129	60	8154	457	> 99999

where $\cos(\beta) = \frac{\lambda_{\min}((\tilde{\mathbf{B}}+\tilde{\mathbf{B}}^T)/2)}{\|\tilde{\mathbf{B}}\|} \geq \frac{\tau}{\Upsilon}\frac{1}{1+C\Delta t p^2/h}$ which, in view of the uniform lower bound on $\frac{\tau}{\Upsilon}$, shows that using the diagonal preconditioner will give results similar to what one expects were the full mass matrix to be used as a preconditioner for $\mathbf{B}$. We display the number of iterations needed for GMRES to converge when solving the matrices $\tilde{\mathbf{B}}$ and $\mathbf{B}$ with $\nu = (1, 1, 1)$ on a cube with 132 elements in Table 3. Observe that preconditioning with the diagonal of the mass matrix proves to be quite effective in reducing iteration count in all cases, even when Δt is relatively large.

3.5. Applicability to Other Types of Basis. The discussion thus far might leave the reader with the (false) impression that our preconditioner is only applicable provided one uses the basis presented in subsection 2.1. This is not the case. The preconditioner is applicable to any choice of basis. Indeed, our preconditioner can be regarded as defining an abstract Additive Schwarz method (ASM) [8, 36] as follows:

The ASM is defined by the following subspace decomposition

$$X = X_I \oplus \bigoplus_{k=1}^{4} X_{F_k} \oplus \bigoplus_{k=1}^{6} X_{E_k} \oplus \bigoplus_{k=1}^{4} X_{V_k},$$

in conjunction with an exact solver on each subspace. Specifically, given a residual $f \in X$, the action of the ASM is defined as follows:

- $u_I \in X_I : (u_I, v_I) = (f, v_I) \quad \forall v_I \in X_I$,
- $u_{F_k} \in X_{F_k} : (u_{F_k}, v_{F_k}) = (f, v_{F_k}) \quad \forall v_{F_k} \in X_{F_k}$,
- $u_{E_k} \in X_{E_k} : (u_{E_k}, v_{E_k}) = (f, v_{E_k}) \quad \forall v_{E_k} \in X_{E_k}$,
- $u_{V_k} \in X_{V_k} : (u_{V_k}, v_{V_k}) = (f, v_{V_k}) \quad \forall v_{V_k} \in X_{V_k}$,

and returns $u := u_I + \sum_{k=1}^{4} u_{F_k} + \sum_{k=1}^{6} u_{E_k} + \sum_{k=1}^{4} u_{V_k}$. This formulation of the preconditioner relies only on the choice of space, and not on the particular basis. The proof that the ASM gives rise to an uniform bound on the condition number follows from the fact that the constants c, C in (2.9) are independent of p [36, Theorem 2.7].

The action of the preconditioner for a general choice of basis consists of first statically condensing out the interior degrees of freedom. Lemma 4.3 states that X_I is L^2 orthogonal to the remaining subspaces:

$$X_I \perp \bigoplus_{k=1}^{4} X_{F_k} \oplus \bigoplus_{k=1}^{6} X_{E_k} \oplus \bigoplus_{k=1}^{4} X_{V_k}$$

which means that one can first reduce the system to the Schur complement matrix. Once the Schur complement is in hand, a change of basis can be applied on the interface to map to the spaces X_{F_k}, X_{E_k} and X_{V_k} corresponding to the preconditioner presented here. Specific details in the 2D setting can be found in [5]. The same approach extends readily to tetrahedral elements considered here; most of the numerical examples of section 3 were computed using the Bernstein basis in conjunction with a change of basis operator.

4. Technical Lemmas.

In this section, we turn to the proof of the technical lemmas which were used in proving Theorem 2.1.

4.1. Orthogonality.

The Duffy transformation [18, §3.2] given by

$$\xi := \frac{2\lambda_2}{1-\lambda_3-\lambda_4} - 1, \qquad \eta := \frac{2\lambda_3}{1-\lambda_4} - 1, \qquad \theta := 2\lambda_4 - 1$$

maps the tetrahedron T onto the cube $\{(\xi,\eta,\theta) : -1 \le \xi,\eta,\theta \le 1\}$. For reference, the edge $E_1 = \{(\xi,\eta,\theta) : -1 \le \xi \le 1, \eta = -1, \theta = -1\}$ and the face $F_1 = \{(\xi,\eta,\theta) : -1 \le \xi,\eta \le 1, \theta = -1\}$.

We begin by establishing the orthogonality properties of the basis functions:

LEMMA 4.1. *The functions* $\{\omega_{ijk}\}, \{\psi_{ij}^{(k)}\}, \{\chi_i^{(k)}\}$ *provide an* L^2*-orthogonal basis for* X_I, X_{F_k}, X_{E_k} *respectively.*

Proof. It suffices to show that

$$(\omega_{i_1j_1k_1}, \omega_{i_2j_2k_2}) \propto \delta_{i_1j_1k_1,i_2j_2k_2}, \quad (\psi_{i_1j_1}^{(1)}, \psi_{i_2j_2}^{(1)}) \propto \delta_{i_1j_1,i_2j_2}, \quad (\chi_{i_1}^{(1)}, \chi_{i_2}^{(1)}) \propto \delta_{i_1,i_2}.$$

Transforming the basis functions using the Duffy transformation gives

$$\begin{aligned}
\omega_{ijk} &= \frac{1-\xi}{2}\frac{1+\xi}{2}P_i^{(2,2)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+2}\frac{1+\eta}{2}P_j^{(2i+5,2)}(\eta) \\
&\qquad \times \left(\frac{1-\theta}{2}\right)^{i+j+3}\frac{1+\theta}{2}P_k^{(2i+2j+8,2)}(\theta), \\
\psi_{ij}^{(1)} &= \frac{1-\xi}{2}\frac{1+\xi}{2}P_i^{(2,2)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+2}\frac{1+\eta}{2}P_j^{(2i+5,2)}(\eta) \\
&\qquad \times \left(\frac{1-\theta}{2}\right)^{i+j+3}\Phi_{p-3-i-j}^{(2i+2j+8)}(\theta), \\
\chi_i^{(1)} &= \frac{1-\xi}{2}\frac{1+\xi}{2}P_i^{(2,2)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+2}\left(\frac{1-\theta}{2}\right)^{i+2}F(\eta,\theta)
\end{aligned}$$

where $F(\eta,\theta)$ is a polynomial in η and θ.

The Jacobian of the Duffy transformation is given by

$$J = \frac{1-\eta}{2}\left(\frac{1-\theta}{2}\right)^2,$$

and, as a consequence, we find

$$\begin{aligned}
\int_T \omega_{i_1j_1k_1}\omega_{i_2j_2k_2}\,dx &= \int_{-1}^{1}\left(\frac{1-\xi}{2}\right)^2\left(\frac{1+\xi}{2}\right)^2 P_{i_1}^{(2,2)}P_{i_2}^{(2,2)}\,d\xi \\
&\quad \times \int_{-1}^{1}\left(\frac{1-\eta}{2}\right)^{i_1+i_2+5}\left(\frac{1+\eta}{2}\right)^2 P_{j_1}^{(2i_1+5,2)}P_{j_2}^{(2i_2+5,2)}\,d\eta \\
&\quad \times \int_{-1}^{1}\left(\frac{1-\theta}{2}\right)^{i_1+i_2+j_1+j_2+8}\left(\frac{1+\theta}{2}\right)^2 P_{k_1}^{(2i_1+2j_1+8,2)}P_{k_2}^{(2i_2+2j_2+8,2)}\,d\theta \\
&= C\delta_{i_1,i_2}\delta_{j_1,j_2}\delta_{k_1,k_2}.
\end{aligned}$$

The result for the edge $\psi_{ij}^{(1)}$ and face $\chi_i^{(1)}$ functions follows the same lines. ∎

The next lemma enumerates the pertinent properties of the function $\Phi_p^{(m)}$ which was used in several places in defining the basis functions:

Lemma 4.2. *For non-negative integers m, q, $\Phi_p^{(m)}$ has the following properties:*

1. $\Phi_q^{(m)}(-1) = 1$,
2. *Weighted norm*

$$I_{m,q} := \int_{-1}^{1} \left(\frac{1-x}{2}\right)^m \left(\Phi_q^{(m)}(x)\right)^2 dx = \frac{2}{(q+1)(m+q+1)}, \tag{4.1}$$

3. *Orthogonality property*

$$\int_{-1}^{1} \left(\frac{1-x}{2}\right)^m \frac{1+x}{2} \Phi_q^{(m)}(x) w(x)\, dx = 0$$

for all $w \in \mathbb{P}_r([-1,1])$ with $r < q$.

Proof. The first property comes from the fact that $P_q^{(m,1)}(-1) = (-1)^q \binom{q+1}{q}$ [1, §22.2.1], and the third property follows straight from the orthogonality property of $P_q^{(m,1)}$. For the second result, relation (22.7.19) in [1] gives us

$$\frac{2q+m+1}{q+m+1} P_q^{(m,0)} - \frac{q+m}{q+m+1} P_{q-1}^{(m,1)} = P_q^{(m,1)}.$$

Equation (4.1) in the case of $q = 0$ trivially holds. Suppose that (4.1) holds in the case of $q-1$, then

$$\begin{aligned}
I_{m,q} &= \frac{1}{(q+1)^2} \int_{-1}^{1} \left(\frac{1-x}{2}\right)^m P_q^{(m,1)}(x) P_q^{(m,1)}(x)\, dx \\
&= \frac{1}{(q+1)^2} \int_{-1}^{1} \left(\frac{1-x}{2}\right)^m \left(\frac{(2q+m+1)^2}{(q+m+1)^2} P_q^{(m,0)}(x) P_q^{(m,0)}(x)\right) dx \\
&\quad + \frac{1}{(q+1)^2} \frac{(q+m)^2}{(q+m+1)^2} q^2 I_{m,q-1} \\
&= \frac{1}{(q+1)^2} \frac{(2q+m+1)^2}{(q+m+1)^2} \frac{2}{2q+m+1} + \frac{1}{(q+1)^2} \frac{(q+m)^2}{(q+m+1)^2} q^2 \frac{2}{q(m+q)} \\
&= \frac{2}{(q+1)(q+m+1)}
\end{aligned}$$ ∎

and the result (4.1) holds by induction.

The above result implies that the interior basis functions are orthogonal to the face/edge/vertex functions:

Lemma 4.3. *Let $X_B = \bigoplus_{k=1}^4 X_{F_k} \oplus \bigoplus_{k=1}^6 X_{E_k} \oplus \bigoplus_{k=1}^4 X_{V_k}$, then the space X can be decomposed as $X = X_I \oplus X_B$ such that $X_I \perp X_B$.*

Proof. Recall Ξ_i, q_i and q from (2.2)–(2.4) respectively, and define $\bar{\chi}_i^{(1)}, \bar{\varphi}_1$ as

(4.2)
$$\begin{aligned}\bar{\chi}_i^{(1)} &:= \lambda_1\lambda_2\Xi_i(\lambda_1,\lambda_2)q_i(\lambda_3,\lambda_4)\\ &= \frac{1-\xi}{2}\frac{1+\xi}{2}P_i^{(2,2)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+2}\Phi_j^{(2i+5)}(\eta)\left(\frac{1-\theta}{2}\right)^{i+j+2}\Phi_{p-2-i-j}^{(2i+2j+6)}(\theta),\\ \bar{\varphi}_1 &:= \lambda_1 q(\lambda_2,\lambda_3,\lambda_4)\\ &= \frac{1-\xi}{2}\Phi_i^{(2)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+1}\Phi_j^{(2i+3)}(\eta)\left(\frac{1-\theta}{2}\right)^{i+j+1}\Phi_{p-1-i-j}^{(2i+2j+4)}(\theta).\end{aligned}$$

By permutation of the barycentric coordinates, it suffices to show that for any interior basis function ω_{lmn} with $0 \le l, m, n, l+m+n \le p-4$, the inner product vanishes

$$\begin{aligned}(\bar{\varphi}_1, \omega_{lmn}) &= 0,\\ (\bar{\chi}_i^{(1)}, \omega_{lmn}) &= 0, \quad i = 0,\dots,p-2,\\ (\psi_{ij}^{(1)}, \omega_{lmn}) &= 0, \quad 0 \le i, j, i+j \le p-3.\end{aligned}$$

Calculating the inner-product for the face functions first:

$$\begin{aligned}(\psi_{ij}^{(1)}, \omega_{lmn}) = &\int_{-1}^{1}\left(\frac{1-\xi}{2}\right)^2\left(\frac{1+\xi}{2}\right)^2 P_i^{(2,2)}(\xi)P_l^{(2,2)}(\xi)\,d\xi\\ &\times\int_{-1}^{1}\left(\frac{1-\eta}{2}\right)^{i+l+5}\left(\frac{1+\eta}{2}\right)^2 P_j^{(2i+5,2)}(\eta)P_m^{(2l+5,2)}(\eta)\,d\eta\\ &\times\int_{-1}^{1}\left(\frac{1-\theta}{2}\right)^{i+l+j+m+8}\left(\frac{1+\theta}{2}\right)\Phi_{p-3-i-j}^{(2i+2j+8)}(\theta)P_n^{(2l+2m+8,2)}(\theta)\,d\theta\\ \propto\ &\delta_{il}\delta_{jm}\int_{-1}^{1}\left(\frac{1-\theta}{2}\right)^{2i+2j+8}\left(\frac{1+\theta}{2}\right)\Phi_{p-3-i-j}^{(2i+2j+8)}(\theta)P_n^{(2l+2m+8,2)}(\theta)\,d\theta.\end{aligned}$$

The inner-product vanishes if $i \ne l, j \ne m$. Assuming otherwise, then we have that $p-3-i-j > n$ as $l+m+n \le p-4$, hence the inner-product is 0 by Lemma 4.2.

For the edges, we have

$$\begin{aligned}(\bar{\chi}_i^{(1)}, \omega_{lmn}) \propto \delta_{il}&\int_{-1}^{1}\left(\frac{1-\eta}{2}\right)^{i+l+5}\frac{1+\eta}{2}P_j^{(2i+5,1)}(\eta)P_m^{(2l+5,2)}(\eta)\,d\eta\\ &\times\int_{-1}^{1}\left(\frac{1-\theta}{2}\right)^{i+j+l+m+7}\frac{1+\theta}{2}P_{p-2-i-j}^{(2i+2j+6,1)}(\theta)P_n^{(2l+2m+8,2)}(\theta)\,d\theta.\end{aligned}$$

The inner product is trivially zero if $i \ne l$ or $m < j$. Assuming otherwise, we have for the θ variable

$$\int_{-1}^{1}\left(\frac{1-\theta}{2}\right)^{2i+2j+6}\frac{1+\theta}{2}\left[\left(\frac{1-\theta}{2}\right)^{1+m-j}P_n^{(2l+2m+8,2)}(\theta)\right]P_{p-2-i-j}^{(2i+2j+6,1)}(\theta)\,d\theta.$$

The above vanishes if

$$1+m-j+n < p-2-i-j$$

which follows from the fact that $l+m+n \le p-4$.

Finally, we have

$$(\bar{\varphi}_1, \omega_{lmn}) \propto \int_{-1}^{1} \left(\frac{1-\xi}{2}\right)^2 \frac{1+\xi}{2} P_i^{(2,1)}(\xi) P_l^{(2,2)}(\xi)\, d\xi$$
$$\int_{-1}^{1} \left(\frac{1-\eta}{2}\right)^{i+l+4} \frac{1+\eta}{2} P_j^{(2i+3,1)}(\eta) P_m^{(2l+5,2)}(\eta)\, d\eta$$
$$\int_{-1}^{1} \left(\frac{1-\theta}{2}\right)^{i+j+l+m+6} \frac{1+\theta}{2} P_k^{(2i+2j+4,1)}(\theta) P_l^{(2l+2m+8,2)}(\theta)\, d\theta.$$

If $i > l$, then there is nothing to prove, otherwise the η integral can be written as

$$\int_{-1}^{1} \left(\frac{1-\eta}{2}\right)^{2i+3} \frac{1+\eta}{2} \left[\left(\frac{1-\eta}{2}\right)^{1+l-i} P_m^{(2l+5,2)}(\eta)\right] P_j^{(2i+3,1)}(\eta)\, d\eta$$

which vanishes if $j > 1+l-i+m$. Finally, assuming otherwise, the θ integral can be written as

$$\int_{-1}^{1} \left(\frac{1-\theta}{2}\right)^{2i+2j+4} \frac{1+\theta}{2} \left[\left(\frac{1-\theta}{2}\right)^{l+m-i-j+2} P_n^{(2l+2m+8,2)}(\theta)\right] P_{p-1-i-j}^{(2i+2j+4,1)}(\theta)\, d\theta.$$

The above quantity vanishes if

$$l+m-i-j+2+n < p-1-i-j$$

which follows from the fact that $l+m+n \le p-4$. ∎

Now we turn to the stability of the subspace decomposition.

4.2. Vertex Contributions. The following lemma corresponds to Lemma 5.4 and 6.1 of [4] and allows us to bound the vertex contribution:

LEMMA 4.4. *The vertex basis functions of degree p satisfy the bound*

$$cp^{-3} \le \|\varphi\| \le Cp^{-3}$$

for constants c, C independent of p.

Proof. Note that

$$\|\varphi_1\| = \|\bar{\varphi}_1/3 + \lambda_1 q(\lambda_3,\lambda_4,\lambda_2)/3 + \lambda_1 q(\lambda_4,\lambda_2,\lambda_3)/3\|$$
$$\le \|\bar{\varphi}_1/3\| + \|\lambda_1 q(\lambda_3,\lambda_4,\lambda_2)/3\| + \|\lambda_1 q(\lambda_4,\lambda_2,\lambda_3)/3\| = \|\bar{\varphi}_1\|$$

where $\bar{\varphi}_1$ is defined in (4.2).

Using Lemma 4.2,

$$\|\bar{\varphi}\|^2 = \int_{-1}^{1} \frac{(1-\xi)^2}{4} \Phi_i^{(2)}\, d\xi \int_{-1}^{1} \left(\frac{1-\eta}{2}\right)^{2i+3} \Phi_j^{(2i+3)}\, d\eta$$
$$\times \int_{-1}^{1} \left(\frac{1-\theta}{2}\right)^{2i+2j+4} \Phi_{p-1-i-j}^{(2i+2j+4)}\, d\theta$$
$$= \frac{8}{(i+1)(i+3)(j+1)(2i+j+4)(p-i-j)(i+j+p+4)} \le Cp^{-6}.$$

For the lower bound, let $0 \leq i, j, k, i+j+k \leq p$ and define

$$\Psi_{ijk} := c_{ijk} P_i^{(0,0)}(\xi) \left(\frac{1-\eta}{2}\right)^i P_j^{(2i+1,0)}(\eta) \left(\frac{1-\theta}{2}\right)^{i+j} P_k^{(2i+2j+2,0)}(\theta), \tag{4.3}$$

where $c_{ijk} = \frac{1}{2}\sqrt{(2i+1)(i+j+1)(2i+2j+2k+3)}$. These functions form an orthonormal basis for X hence φ can be written in the form $\varphi = \sum_{i+j+k \leq p} u_{ijk} \Psi_{ijk}$ where u_{ijk} are the appropriate coefficients and $\|\varphi\|^2 = \sum_{i+j+k \leq p} u_{ijk}^2$. It suffices to prove the inequality in the case of φ_1. Cauchy-Schwarz gives

$$\begin{aligned} 1 = |\varphi(-1,-1,-1)|^2 &= \left(\sum_{i+j+k \leq p} (-1)^{i+j+k} c_{ijk} u_{ijk} \right)^2 \\ &\leq \sum_{i+j+k \leq p} u_{ij}^2 \sum_{i+j+k \leq p} c_{ijk}^2 = \frac{(p+1)^2(p+2)^2(p+3)^2}{48} \|\varphi\|^2 . \quad \square \end{aligned}$$

We now proceed to the edge contributions.

4.3. Edge contributions.

The following lemma bounds the contribution on an edge:

LEMMA 4.5. *Let $u \in X$ be such that u vanishes at the vertices of T. Let γ be an arbitrary edge of T and let $U \in X_{E_\gamma}$ such that $U|_\gamma = u|_\gamma$. Then there exists a constant C independent of p such that*

$$\|U\| \leq C\|u\| . \tag{4.4}$$

Proof. Without loss of generality, we assume that $\gamma := E_1$. Let $U = \sum_{i=0}^{p-2} w_i \chi_i^{(1)}$ where the coefficients w_i are chosen such that $U|_\gamma = u|_\gamma$. It is more convenient to work with the function $\bar{\chi}_i^{(1)}$ defined in (4.2). Observe that $\bar{\chi}_i^{(1)}|_{E_1} = \chi_i^{(1)}|_{E_1}$, and $(\bar{\chi}_i^{(1)}, \bar{\chi}_j^{(1)}) \propto \delta_{ij}$. Let $\bar{U} = \sum_{i=0}^{p-2} w_i \bar{\chi}_i^{(1)}$, then $\bar{U} = U$ on edge γ and $\|U\| \leq \|\bar{U}\|$ as

$$\begin{aligned} \left\|\chi_i^{(1)}\right\| &= \left\|\bar{\chi}_i^{(1)}/2 + \lambda_1\lambda_2 \Xi_i(\lambda_1,\lambda_2) q_j(\lambda_4,\lambda_3)/2\right\| \\ &\leq \left\|\bar{\chi}_i^{(1)}/2\right\| + \left\|\lambda_1\lambda_2 \Xi_i(\lambda_1,\lambda_2) q_j(\lambda_4,\lambda_3)/2\right\| = \left\|\bar{\chi}_i^{(1)}\right\|, \end{aligned}$$

thus it suffices to show that $\|\bar{U}\| \leq C\|u\|$.

To this end, recall the orthonormal basis Ψ_{ijk} defined in (4.3) and let $u = \sum_{i+j+k \leq p} u_{ijk} \Psi_{ijk}$ and

$$f := u|_\gamma = \sum_{i=0}^{p} v_i P_i^{(0,0)}(x)$$

where

$$v_i := \sum_{j=0}^{p-i} \sum_{k=0}^{p-i-j} \frac{(-1)^{j+k}}{2} u_{ijk} \sqrt{(2i+1)(i+j+1)(2i+2j+2k+3)}, \tag{4.5}$$

Furthermore, since u vanishes at the vertices of T, then $f(\pm 1) = 0$ thus

$$\sum_{i=0,\text{even}}^{p} v_i = 0, \qquad \sum_{i=1,\text{odd}}^{p} v_i = 0. \tag{4.6}$$

Consequently, we can rewrite $f = \sum_{i=2}^{p}(P_i^{(0,0)} - P_{i-2}^{(0,0)})S_i$ where

$$S_i = v_i + v_{i+2} + \cdots + \begin{cases} v_p \\ v_{p-1} \end{cases} = \begin{cases} -v_0 - \cdots - v_{i-2} & \text{if } i \text{ even} \\ -v_1 - \cdots - v_{i-2} & \text{else} \end{cases}$$

depending on the parity.

Turning to the coefficients w_i, we must have on edge γ

$$\bar{U}|_\gamma = \frac{1-\xi}{2}\frac{1+\xi}{2}\sum_{i=0}^{p-2} w_i P_i^{(2,2)}(\xi) = \sum_{i=2}^{p}(P_i^{(0,0)} - P_{i-2}^{(0,0)})S_i$$

Recall the following identity from Lemma 6.6 of [4]

$$-\frac{1-x^2}{2(n-1)}\left(\frac{(n+1)(n+2)}{2n}P_{n-2}^{(2,2)} - \frac{n-1}{2}P_{n-4}^{(2,2)}\right) = P_n^{(0,0)} - P_{n-2}^{(0,0)}, \qquad n \geq 2$$

where P_{n-4} is understood to be 0 for $n < 4$, then we have

$$\sum_{i=0}^{p-2} w_i P_i^{(2,2)} = \sum_{i=2}^{p}\left(-\frac{(i+1)(i+2)}{i(i-1)}P_{i-2}^{(2,2)} + P_{i-4}^{(2,2)}\right)S_i$$

and we deduce by matching coefficients that

$$\begin{aligned} w_i &= S_{i+4} - \frac{(i+3)(i+4)}{(i+1)(i+2)}S_{i+2} \\ &= -v_{i+2} - \frac{2(5+2i)}{(i+1)(i+2)}S_{i+2}. \end{aligned} \tag{4.7}$$

With (4.7) in hand, we can now analyze $\|\bar{U}\|$ and $\|u\|$. The Cauchy-Schwarz inequality applied to (4.5) gives

$$\begin{aligned} v_i^2 &\leq \sum_{j=0}^{p-i}\sum_{k=0}^{p-i-j} u_{ijk}^2 \sum_{j=0}^{p-i}\sum_{k=0}^{p-i-j} \frac{(2i+1)(i+j+1)(2i+2j+2k+3)}{4} \\ &= \frac{1}{16}(2i+1)(i-p-2)(i-p-1)(i+p+2)(i+p+3)\sum_{j=0}^{p-i}\sum_{k=0}^{p-i-j} u_{ijk}^2, \end{aligned}$$

hence, rearranging and summing over the index i, we have a lower bound for $\|u\|$

$$\begin{aligned} &\sum_{i=0}^{p} \frac{16v_i^2}{(2i+1)(i-p-2)(i-p-1)(i+p+2)(i+p+3)} \\ &\qquad \approx \sum_{i=0}^{p} \frac{v_i^2}{(i+1)(i-p-1)^2(i+p+1)^2} \leq \|u\|^2. \end{aligned} \tag{4.8}$$

Using Lemma 4.2, the fact that $j = \lfloor \frac{p-i-2}{2} \rfloor$, and Cauchy-Schwarz on (4.7) gives

$$\begin{aligned}
\left\|\bar{U}\right\|^2 &= \sum_{i=0}^{p-2} \frac{2(i+1)(i+2)w_i^2}{(i+3)(i+4)(2i+5)} \frac{2}{(j+1)(2i+j+6)} \frac{2}{(p-i-j-1)(i+j+p+5)} \\
&\approx \sum_{i=0}^{p-2} \frac{w_i^2}{(i+1)} \frac{1}{(p-i+1)(p+i+1)} \frac{1}{(p-i+1)(i+p+1)} \\
&\le C \left(\sum_{i=0}^{p-2} \frac{v_{i+2}^2}{(i+1)(p-i+1)^2(p+i+1)^2} + \frac{S_{i+2}^2}{(i+1)^3(p-i+1)^2(p+i+1)^2} \right).
\end{aligned}$$

The first term is bounded easily by using (4.8)

$$\sum_{i=0}^{p-2} \frac{v_{i+2}^2}{(i+1)(p-i+1)^2(p+i+1)^2} \le C \sum_{i=0}^{p} \frac{v_i^2}{(i+1)(i-p-1)^2(i+p+1)^2} \le C\|u\|^2.$$

Hence, the theorem follows if there exists a constant C independent of p such that

$$\sum_{i=0}^{p-2} \frac{S_{i+2}^2}{(i+1)^3(p-i+1)^2(p+i+1)^2} \le C \sum_{i=0}^{p} \frac{v_i^2}{(i+1)(i-p-1)^2(i+p+1)^2},$$

but this follows by applying Lemma 4.10 with $j = 2$. ∎

4.4. Face contributions. Finally, it remains to show that the face contributions are bounded. Let F be an arbitrary face of T, and let S be a subset of the remaining faces of T. We remark that $S \cup F$ need not necessarily coincide with the set of *all* faces of T. Let $Y_F := \{u \in X : u = 0 \text{ on all the edges of } F\}$, and define the operator $\mathcal{E}_{S,F} : Y_F \mapsto Y_F$ by

$$\mathcal{E}_{S,F} u := \underset{\substack{v|_F = u|_F \\ v|_S = 0 \\ v \in Y_F}}{\operatorname{argmin}} \|v\|^2. \tag{4.9}$$

Existence to the minimization problem is trivial, while uniqueness comes from the strict convexity of the squared L^2 norm. Clearly,

$$\left\|\mathcal{E}_{S\setminus F',F} u\right\| \le \left\|\mathcal{E}_{S,F} u\right\|, \quad \forall F' \subset S$$

since $\mathcal{E}_{S,F} u = u$ on F and also vanishes on $S \setminus F'$. The proof that the converse inequality is also independent of p is less obvious:

LEMMA 4.6. *Let F be an arbitrary face of T, and let S be a subset of the remaining faces of T. There exists a constant C independent of p such that*

$$\left\|\mathcal{E}_{S,F} u\right\| \le C \left\|\mathcal{E}_{S\setminus F',F} u\right\|, \qquad \forall u \in Y_F,$$

for all $F' \subset S$.

Before giving the proof, we note the following consequence of Lemma 4.6 which was used in the proof of Theorem 2.1:

COROLLARY 4.7. *Let F_i be any face of T and $u \in Y_{F_i}$, then there exists a polynomial $U \in X_{F_i}$ such that $U|_{F_i} = u|_{F_i}$ and*

$$\|U\| \le C\|u\|$$

where C is independent of p.

Proof. Choosing $S = \partial T \setminus F_i$, $F' = S$, and let $U = \mathcal{E}_{S,F_i} u$. Clearly, $U \in X_{F_i}$ as U vanishes on S the three remaining faces. Furthermore, Lemma 4.6 gives the bound

$$\|U\| = \left\|\mathcal{E}_{S,F_i} u\right\| \le C\left\|\mathcal{E}_{S\setminus F',F_i} u\right\| \le C\|u\|\,. \qquad \square$$

All that remains is to prove Lemma 4.6; to this end, for $l, m, n \in \{0, 1\}$ define the polynomials

$$\begin{aligned}\zeta_{ij}^{(l,m,n)} = &\left(\frac{1-\xi}{2}\right)^m \left(\frac{1+\xi}{2}\right)^n P_i^{(2m,2n)}(\xi) \left(\frac{1-\eta}{2}\right)^{i+m+n} \left(\frac{1+\eta}{2}\right)^l \\ &\times P_j^{(2i+2m+2n+1,2l)}(\eta) \left(\frac{1-\theta}{2}\right)^{j+i+m+n+l} \Phi_{p-i-j-m-n-l}^{(2(j+i+m+n+l)+2)}(\theta)\end{aligned} \tag{4.10}$$

with $0 \le i, j, i + j \le p - l - m - n$.

LEMMA 4.8. *The following properties hold:*
1. $\zeta_{ij}^{(l,m,n)} \in X$,
2. $\zeta_{ij}^{(1,1,1)}$ *vanishes on* $\{\xi = \pm 1, \eta = 1\}$, $\zeta_{ij}^{(0,1,1)}$ *vanishes on* $\{\xi = \pm 1\}$ *etc.,*
3. $\zeta_{ij}^{(1,1,1)} = \psi_{ij}^{(1)}$, *our face basis functions,*
4. $\{\zeta_{ij}^{(l,m,n)}\}$ *is orthogonal on T for a fixed l, m, n,*
5. $\{\zeta_{ij}^{(l,m,n)}|_{F_1}\}$ *spans* $\mathbb{P}_p(F_1) \cap H_0^1(F_1)$.

Proof. The first three statements can be deduced by inspection. For the orthogonality property, we note that

$$(\zeta_{i_1 j_1}^{(l,m,n)}, \zeta_{i_2 j_2}^{(l,m,n)}) \propto F(\theta) \int_{-1}^{1} \left(\frac{1-\xi}{2}\right)^{2m} \left(\frac{1+\xi}{2}\right)^{2n} P_{i_1}^{(2m,2n)} P_{i_2}^{(2m,2n)}\, d\xi$$

$$\times \int_{-1}^{1} \left(\frac{1-\eta}{2}\right)^{i_1+i_2+2m+2n+1} \left(\frac{1+\eta}{2}\right)^{2l} P_{j_1}^{(2i_1+2m+2n+1,2l)} P_{j_2}^{(2i_2+2m+2n+1,2l)}\, d\eta.$$

The quantity vanishes if $i_1 \ne i_2$ or $j_1 \ne j_2$.

The last statement follows from linear independence, and recognizing that the restriction of the 3D Duffy transformation onto F_1 reduces to the 2D Duffy transformation. $\square$

The following lemma gives an explicit expression for the operator $\mathcal{E}_{S,F}$ defined in (4.9):

LEMMA 4.9. *Let $u \in Y_{F_1}$ then*

$$\mathcal{E}_{S,F_1} u = \sum_{i+j \le p-l-m-n} u_{ij}^{(l,m,n)} \zeta_{ij}^{(l,m,n)} \tag{4.11}$$

where $u_{ij}^{(l,m,n)}$ are determined by the condition

$$\sum_{i+j \le p-l-m-n} u_{ij}^{(l,m,n)} \zeta_{ij}^{(l,m,n)}(\xi, \eta, -1) = u(\xi, \eta, -1) \tag{4.12}$$

and the coefficients l, m, n are given by one of the following conditions depending on S:

1. $S = \{\xi = -1\} \cup \{\xi = 1\} \cup \{\eta = -1\}$, $m = n = l = 1$.
2. $S = \{\xi = -1\} \cup \{\xi = 1\}$, $m = n = 1, l = 0$.
3. $S = \{\xi = -1\} \cup \{\eta = -1\}$, $m = 1, n = 0, l = 1$.
4. $S = \{\xi = 1\} \cup \{\eta = -1\}$, $m = 0, n = l = 1$.
5. $S = \{\xi = -1\}$, $m = 1, n = l = 0$.
6. $S = \{\eta = -1\}$, $m = n = 0, l = 1$.
7. $S = \{\xi = 1\}$, $m = 0, n = 1, l = 0$.
8. $S = \emptyset$, $m = n = l = 0$.

Proof. Clearly, the coefficients $u_{ij}^{(l,m,n)}$ are uniquely defined by (4.12) thanks to properties 4 and 5 of Lemma 4.8. For the sake of notation, we will drop the (l, m, n) notation in the remainder of the proof. It suffices to show that the right hand side of (4.11) solves the minimization problem (4.9).

By statement 4 of Lemma 4.8, and statement 2 of Lemma 4.2, we can calculate

$$\left\| \sum_{i+j\le p-l-m-n} u_{ij}\zeta_{ij} \right\|^2 = \sum_{i+j\le p-l-m-n} u_{ij}^2 \left\|\zeta_{ij}\right\|^2 \tag{4.13}$$
$$= \sum_{i+j\le p-l-m-n} u_{ij}^2 \mu_i \nu_j \frac{2}{(p-i-j-m-n-l+1)(p+i+j+m+n+l+3)}$$

where

$$\mu_i = \int \left(\frac{1-x}{2}\right)^{2m} \left(\frac{1+x}{2}\right)^{2n} (P_i^{(2m,2n)})^2 \, dx$$
$$\nu_j = \int \left(\frac{1-x}{2}\right)^{2i+2m+2n+1} \left(\frac{1+x}{2}\right)^{2l} (P_j^{(2i+2m+2n+1,2l)})^2 \, dx.$$

We will show below that $\left\|\mathcal{E}_{S,F_1} u\right\|^2$ equals the above quantity (4.13).

For $i + j + k \le p - l - m - n$, let

$$\Psi_{ijk} := \left(\frac{1-\xi}{2}\right)^m \left(\frac{1+\xi}{2}\right)^n P_i^{(2m,2n)}(\xi) \left(\frac{1-\eta}{2}\right)^{i+m+n} \left(\frac{1+\eta}{2}\right)^l$$
$$\times P_j^{(2i+2m+2n+1,2l)}(\eta) \left(\frac{1-\theta}{2}\right)^{i+m+n+l+j} P_k^{(2(i+m+n+l+j)+2,0)}(\theta).$$

By construction, Ψ_{ijk} vanish on S and are orthogonal to each other, hence there exists coefficients $\widetilde{u}_{ijk}$ such that $\mathcal{E}_{S,F_1} u = \sum_{i+j+k\le p-m-n-l} \widetilde{u}_{ijk}\Psi_{ijk}$ with

$$\left\|\mathcal{E}_{S,F_1} u\right\|^2 = \sum_{i+j+k\le p-m-n-l} \widetilde{u}_{ijk}^2 \mu_i \nu_j \rho_k$$

where

$$\rho_k = \int \left(\frac{1-x}{2}\right)^{2(i+m+n+l+j)+2} (P_k^{(2(i+m+n+l+j)+2,0)})^2 \, dx.$$

We now turn to the relationship between u_{ij} and $\widetilde{u}_{ijk}$. First, note that $\zeta_{ij}|_{F_1} = \Psi_{ijk}|_{F_1}$ hence in order to satisfy the constraint on F_1, we must have $\sum u_{ij}\zeta_{ij}|_{F_1} = \sum \widetilde{u}_{ijk}\Psi_{ijk}|_{F_1}$ and thus

$$u_{ij} = \sum_{k=0}^{p-i-j-m-n-l} \widetilde{u}_{ijk} P_k^{(2(i+m+n+l+j)+2,0)}(-1) = \sum_{k=0}^{p-i-j-m-n-l} (-1)^k \widetilde{u}_{ijk}. \tag{4.14}$$

By Cauchy-Schwarz inequality, we have that

$$u_{ij}^2 \leq \sum_{k=0}^{p-i-j-m-n-l} \widetilde{u}_{ijk}^2 \rho_k \sum_{k=0}^{p-i-j-m-n-l} \rho_k^{-1} \tag{4.15}$$

which implies a lower bound for the norm of the extension in terms of u_{ij}

$$\begin{aligned} \left\|\mathcal{E}_{S,F_1} u\right\|^2 &= \sum_{i+j+k \leq p-m-n-l} \widetilde{u}_{ijk}^2 \mu_i \nu_j \rho_k \\ &\geq \sum_{i=0}^{p-m-n-l} \mu_i \sum_{j=0}^{p-m-n-l-i} \nu_j \frac{u_{ij}^2}{\sum_{k=0}^{p-i-j-m-n-l} \rho_k^{-1}}. \end{aligned} \tag{4.16}$$

In fact, equality can be achieved in (4.15) if we let

$$\widetilde{u}_{ijk} = (-1)^k \rho_k^{-1} \left(\frac{u_{ij}}{\sum_{k=0}^{p-i-j-m-n-l} \rho_k^{-1}} \right).$$

One can verify that with this choice of coefficients that (4.14) is still satisfied. As $\rho_k = \frac{2}{2(i+j+l+m+n)+2k+3}$, thus

$$\sum_{k=0}^{p-i-j-m-n-l} \rho_k^{-1} = \frac{1}{2}(p-i-j-l-m-n+1)(i+j+l+m+n+p+3).$$

Comparing (4.16) with (4.13), we see that they are indeed equal. ∎

Finally we are in a position to give the proof of Lemma 4.6:

Proof. We first prove the case where F' consists of a single face. Without loss of generality, we can assume that $F = F_1 = \{\theta = -1\}$ the reference face, and $F' = \{\eta = -1\}$. There are three cases corresponding to $S \setminus F'$ consisting of the empty set, a single face or two faces:
Case 1. If $S = F'$, we choose $m = n = 0$.
Case 2. If $S \setminus F'$ is a single face, we choose $m = 0, n = 1$ or $m = 1, n = 0$.
Case 3. If $S \setminus F'$ consists of the two remaining faces, we choose $m = n = 1$.

Let $\alpha, \beta \in X$ be

$$\alpha := \sum_{i+j \leq p-1-m-n} \alpha_{ij} \zeta_{ij}^{(1,m,n)}, \qquad \beta := \sum_{i+j \leq p-m-n} \beta_{ij} \zeta_{ij}^{(0,m,n)}$$

with coefficients α_{ij}, β_{ij} such that α and β coincides with u on face F_1 (i.e. $u|_{F_1} = \alpha(\xi,\eta,-1) = \beta(\xi,\eta,-1)$). Lemma 4.9 implies that

$$\alpha = \mathcal{E}_{S,F_1} u, \qquad \beta = \mathcal{E}_{S \setminus F',F_1} u,$$

and it suffices to show that there exists a C independent of p such that $\|\alpha\| \le C\|\beta\|$.

Using orthogonality of the basis functions and Lemma 4.2 gives

$$\begin{aligned}\|\alpha\|^2 &= \sum_{i+j\le p-1-m-n} \frac{2(i+2m)!(i+2n)!\alpha_{ij}^2}{i!(2i+2m+2n+1)(i+2(m+n))!} \\ &\quad\times \frac{(j+1)(j+2)}{(i+j+m+n+2)(2i+j+2m+2n+3)(2i+j+2(m+n+1))} \\ &\quad\times \frac{2}{(p-i-j-m-n)(i+j+m+n+p+4)} \\ &\approx \sum_{i+j\le p-1-m-n} \frac{2(i+2m)!(i+2n)!\alpha_{ij}^2}{i!(2i+2m+2n+1)(i+2(m+n))!} \\ &\quad\times \frac{(j+1)^2}{(i+j+1)^3}\frac{1}{(p-i-j)(i+j+p)}\end{aligned} \tag{4.17}$$

and

$$\begin{aligned}\|\beta\|^2 &= \sum_{i+j\le p-m-n} \frac{2(i+2m)!(i+2n)!\beta_{ij}^2}{i!(2i+2m+2n+1)(i+2(m+n))!} \\ &\quad\times \frac{1}{i+j+m+n+1}\frac{2}{(p-i-j-m-n+1)(i+j+m+n+p+3)} \\ &\approx \sum_{i+j\le p-m-n} \frac{2(i+2m)!(i+2n)!\beta_{ij}^2}{i!(2i+2m+2n+1)(i+2(m+n))!} \\ &\quad\times \frac{1}{i+j+1}\frac{1}{(p-i-j+1)(i+j+p)}.\end{aligned} \tag{4.18}$$

We thus have to show for all $0 \le i \le p-m-n-1$ that

$$\begin{aligned}&\sum_{j=0}^{p-1-m-n-i} \frac{(j+1)^2\alpha_{ij}^2}{(i+j+1)^3}\frac{1}{(p-i-j)(i+j+p)} \\ &\qquad\qquad \le C\sum_{j=0}^{p-m-n-i} \frac{\beta_{ij}^2}{i+j+1}\frac{1}{(p-i-j+1)(i+j+p)}.\end{aligned} \tag{4.19}$$

Now, we turn to the relationship between the coefficients α_{ij} and β_{ij}. First, note that since $u \in Y_{F_1}$, it vanishes on the edges of F_1; in particular $u|_{F_1\cap\{\eta=-1\}} = 0$. We have $\alpha|_{F_1\cap\{\eta=-1\}} = 0$ as $\zeta_{ij}^{(1,m,n)}$ vanishes on $\eta=-1$, but the basis functions of β does not vanishes trivially on $\eta=-1$. We see that

$$\begin{aligned}\beta|_{F_1\cap\{\eta=-1\}} &= \sum_{i+j\le p-m-n}\left(\frac{1-\xi}{2}\right)^m\left(\frac{1+\xi}{2}\right)^n P_i^{(2m,2n)}(\xi)(-1)^j\beta_{ij} \\ &= \sum_{i=0}^{p-m-n}\left(\frac{1-\xi}{2}\right)^m\left(\frac{1+\xi}{2}\right)^n P_i^{(2m,2n)}(\xi)\sum_{j=0}^{p-m-n-i}(-1)^j\beta_{ij},\end{aligned}$$

hence by linear independence,

$$\sum_{j=0}^{p-m-n-i}(-1)^j\beta_{ij} = 0 \tag{4.20}$$

in order for $\beta|_{F_1\cap\{\eta=-1\}}$ to vanish.

Now returning to the face F_1, let $\gamma = 2i+2m+2n+1$, then

$$\alpha|_{F_1} = \sum_{i=0}^{p-1-m-n}\left(\frac{1-\xi}{2}\right)^m\left(\frac{1+\xi}{2}\right)^n P_i^{(2m,2n)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+m+n} \times \sum_{j=0}^{p-1-m-n-i}\left(\frac{1+\eta}{2}\right)P_j^{(\gamma,2)}(\eta)\alpha_{ij}$$

By (4.20), $\beta_{p-m-n,0}=0$ hence

$$\begin{aligned}\beta|_{F_1} &= \sum_{i=0}^{p-m-n}\left(\frac{1-\xi}{2}\right)^m\left(\frac{1+\xi}{2}\right)^n P_i^{(2m,2n)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+m+n}\sum_{j=0}^{p-m-n-i}P_j^{(\gamma,0)}(\eta)\beta_{ij}\\ &= \sum_{i=0}^{p-m-n-1}\left(\frac{1-\xi}{2}\right)^m\left(\frac{1+\xi}{2}\right)^n P_i^{(2m,2n)}(\xi)\left(\frac{1-\eta}{2}\right)^{i+m+n}\sum_{j=0}^{p-m-n-i}P_j^{(\gamma,0)}(\eta)\beta_{ij}.\end{aligned}$$

As $\alpha|_{F_1}=\beta|_{F_1}$, then we must have that for a fixed $0\le i\le p-1-m-n$

$$\sum_{j=0}^{p-m-n-i-1}\alpha_{ij}\left(\frac{1+\eta}{2}\right)P_j^{(\gamma,2)}(\eta) = \sum_{j=0}^{p-m-n-i}\beta_{ij}P_j^{(\gamma,0)}(\eta).$$

By telescoping the sum, we have

$$\sum_{j=0}^{p-m-n-i}\beta_{ij}P_j^{(\gamma,0)}(\eta) = \sum_{j=0}^{p-m-n-i}S_{ij}(P_{j+1}^{(\gamma,0)}(\eta)+P_j^{(\gamma,0)}(\eta)) \tag{4.21}$$

where $S_{ij}=\sum_{k=0}^{j}(-1)^{k+j}\beta_{ik}$ with $S_{i,p-m-n-i}=0$ due to (4.20).

Combining (22.7.16) and (22.7.19) of [1] gives the following relation

$$P_{j+1}^{(\gamma,0)}(x)+P_j^{(\gamma,0)}(x) = \frac{x+1}{2}\left(\frac{(\gamma+j)}{j+1}P_{j-1}^{(\gamma,2)}(x)+\frac{\gamma+j+2}{j+1}P_j^{(\gamma,2)}(x)\right) \tag{4.22}$$

for non-negative j where we assume that $P_{-1}^{(\gamma,2)}=0$. Hence, substituting (4.22) into (4.21), we have

$$\sum_{j=0}^{p-m-n-i}\beta_{ij}P_j^{(\gamma,0)}(\eta) = \sum_{j=0}^{p-m-n-i}S_{ij}\frac{\eta+1}{2}\left(\frac{(\gamma+j)}{j+1}P_{j-1}^{(\gamma,2)}(\eta)+\frac{\gamma+j+2}{j+1}P_j^{(\gamma,2)}(\eta)\right).$$

Matching coefficients, we have that

$$\alpha_{ij} = \frac{\gamma+j+2}{j+1}S_{ij}+\frac{\gamma+j+1}{j+2}S_{i,j+1} = \frac{\gamma+j+1}{j+2}\beta_{i,j+1}+\frac{\gamma+2j+3}{(j+1)(j+2)}S_{ij}.$$

Using the inequality $(a+b)^2\le 2a^2+2b^2$, we have that

$$\alpha_{ij}^2 \le 2\left(\frac{\gamma+j+1}{j+2}\right)^2\beta_{i,j+1}^2+2\left(\frac{\gamma+2j+3}{(j+1)(j+2)}\right)^2S_{ij}^2.$$

Inserting the above into (4.19), it suffices to show that there exists a constant C independent of p and i such that

$$\sum_{j=0}^{p-1-m-n-i} \frac{(j+1)^2\left(\frac{\gamma+j+1}{j+2}\right)^2}{(i+j+1)^3} \frac{\beta_{i,j+1}^2}{(p-i-j)(i+j+p)} \leq C \sum_{j=0}^{p-m-n-i} \frac{\beta_{ij}^2}{i+j+1} \frac{1}{(p-i-j+1)(i+j+p)}.$$

and

$$\sum_{j=0}^{p-1-m-n-i} \frac{(j+1)^2\left(\frac{\gamma+2j+3}{(j+1)(j+2)}\right)^2}{(i+j+1)^3} \frac{S_{ij}^2}{(p-i-j)(i+j+p)} \leq C \sum_{j=0}^{p-m-n-i} \frac{\beta_{ij}^2}{i+j+1} \frac{1}{(p-i-j+1)(i+j+p)}.$$

For the first expression, we note that $\gamma+j+1 \approx i+j+1$ hence the inequality follows trivially. As for the second expression, we note that

$$\frac{\gamma+2j+3}{(j+1)(j+2)} \approx \frac{i+j+1}{(j+1)^2}$$

Hence, we wish to show that

$$\sum_{j=0}^{p-1-m-n-i} \frac{S_{ij}^2}{(j+1)^2(i+j+1)} \frac{1}{(p-i-j)(i+j+p)} \leq C \sum_{j=0}^{p-m-n-i} \frac{\beta_{ij}^2}{i+j+1} \frac{1}{(p-i-j+1)(i+j+p)}.$$

By Corollary 4.11, there exists a C independent of p and i, and we are done with the case of F' consisting of a single face.

In the case where F' consists of two or three faces, we can simply bootstrap the argument. For example, if $F' = F_1' \cup F_2'$ where F_1', F_2' are two distinct faces, then

$$\left\|\mathcal{E}_{S,F}u\right\| \leq C\left\|\mathcal{E}_{S\setminus F_1',F}u\right\| \leq C\left\|\mathcal{E}_{S\setminus(F_1'\cup F_2'),F}u\right\| = C\left\|\mathcal{E}_{S\setminus F',F}u\right\|. \qquad \square$$

4.5. Hardy Inequalities. It remains to prove the Hardy inequalities used.

LEMMA 4.10. *Let $\{v_i\}_{i=0}^p \in \mathbb{R}$ satisfy*

$$\sum_{i=0}^{p} v_i = 0, \tag{4.23}$$

then for j a positive integer, there exists a constant $C(j)$ independent of p such that

$$\sum_{i=0}^{p} \frac{S_i^2}{(i+1)^3(i+p+1)^j(p-i+1)^j} \leq C\sum_{i=0}^{p} \frac{v_i^2}{(i+1)(i+p+1)^j(p-i+1)^j}$$

where $S_i = \sum_{k=0}^{i} v_k$.

Proof. By (4.23), we have that $S_i = -\sum_{k=i+1}^{p} v_k$, our inequality follows if

$$\sum_{i=0}^{p/2} \frac{\left(\sum_{k=0}^{i} v_k\right)^2}{(i+1)^3(i+p+1)^j(p-i+1)^j} \leq C \sum_{i=0}^{p/2} \frac{v_i^2}{(i+1)(i+p+1)^j(p-i+1)^j} \tag{4.24}$$

and

$$\sum_{i=p/2+1}^{p} \frac{\left(-\sum_{k=i+1}^{p} v_k\right)^2}{(i+1)^3(i+p+1)^j(p-i+1)^j} \leq C \sum_{i=p/2+1}^{p} \frac{v_i^2}{(i+1)(i+p+1)^j(p-i+1)^j} \tag{4.25}$$

both hold with the constant C independent of p.

Hardy's inequality for weighted sums states that for non-negative a_k, b_i, c_i,

$$\sum_{i=0}^{\infty} \left(\sum_{k=0}^{i} a_k\right)^2 b_i \leq C \sum_{i=0}^{\infty} a_i^2 c_i \tag{4.26}$$

with $C \leq 2\sqrt{2}A$ where $A := \sup_{n\in\mathbb{N}} \left(\sum_{i=n}^{\infty} b_i\right)^{1/2} \left(\sum_{i=0}^{n} c_i^{-1}\right)^{1/2} < \infty$ [19, p. 57]. For (4.24) our result follows if we set $a_i = |v_i|$, $b_i^{-1} = (i+1)^3(i+p+1)^j(p-i+1)^j$ and $c_i^{-1} = (i+1)(i+p+1)^j(p-i+1)^j$ for $i = 0, \ldots, p/2$, and let $a_i = 0, b_i = 0, c_i = 1$ for $i > p/2$. It remains to show that A does not grow with p.

We note that

$$\sum_{i=0}^{n} c_i^{-1} \leq p^{2j} \sum_{i=0}^{n} (i+1) \approx n^2 p^{2j}.$$

Furthermore, the supremum can be reduced to over the interval $n \in [0, p/2]$ due to the padding of zeros, hence

$$\begin{aligned}
A^2 &\approx \sup_{n\in[0,p/2]} n^2 p^{2j} \sum_{i=n}^{p/2} \frac{1}{(i+1)^3(i+p+1)^j(p-i+1)^j} \\
&\leq \sup_{n\in[0,p/2]} n^2 p^{2j} \int_n^{p/2} \frac{1}{(x+1)^3(p-p/2+1)^j p^j}\, dx \\
&\approx \sup_{n\in[0,p/2]} n^2 \left(\frac{1}{2(n+1)^2} - \frac{2}{(p+2)^2}\right) < \infty.
\end{aligned}$$

For (4.25), we first transform the sum such that the index starts at 0 by mapping the indices $i \to p-i, k \to p-k$

$$\sum_{i=0}^{p/2-1} \frac{\left(-\sum_{k=0}^{i-1} v_{p-k}\right)^2}{(p-i+1)^3(2p-i+1)^j(i+1)^j} \leq C \sum_{i=0}^{p/2-1} \frac{v_{p-i}^2}{(p-i+1)(2p-i+1)^j(i+1)^j}.$$

Our result follows if we set $a_i = |v_{p-i}|, b_i^{-1} = (p-i+1)^3(2p-i+1)^j(i+1)^j, c_i^{-1} = (p-i+1)(2p-i+1)^j(i+1)^j$ for $i = 0, \ldots, p/2-1$, and let $a_i = 0, b_i = 0, c_i = 1$ for $i \geq p/2$. It remains to show that A does with not grow with p.

Proceeding similarly as before, note that $\sum_{i=0}^{n} c_i^{-1} \leq (2p)^{j+1} \sum_{i=0}^{n} (i+1)^j \approx p^{j+1} n^{j+1}$. The supremum can be reduced to over the interval $n \in [0, p/2-1]$ as before. Calculating, we have

$$
\begin{aligned}
A^2 &\approx \sup_{n\in[0,p/2-1]} n^{j+1} p^{j+1} \sum_{i=n}^{p/2-1} \frac{1}{(p-i+1)^3(2p-i+1)^j(i+1)^j} \\
&\leq \sup_{n\in[0,p/2-1]} n^{j+1} p \int_n^{p/2} \frac{1}{(p-p/2+1)^3(x+1)^j}\, dx \\
&\approx \sup_{n\in[0,p/2-1]} \frac{n^{j+1}}{p^2} \begin{cases} \frac{2(n+1)(p+2)^j - 2^j(p+2)(n+1)^j}{2(j-1)(n+1)^j(p+2)^j} & j > 1 \\ \log\left(\frac{p}{2n}\right) & j = 1 \end{cases} \\
&< \infty. \qquad \square
\end{aligned}
$$

The case $j = 1$ corresponds to Lemma 6.5 of [4] in which it was stated (but not proved explicitly) that the constant $C(1)$ is independent of p. Lemma 4.10 deals with the general case $j \in \mathbb{N}$ and in addition proves explicitly that $C(j)$ is independent of p.

The following Hardy inequality is required for the face extension inequalities:

Corollary 4.11. *Let* $\{v_i\}_{i=0}^{p-k} \in \mathbb{R}$ *where* k *is an integer* $1 \leq k \leq p$, *and* $S_i = \sum_{j=0}^{i} (-1)^j v_j$, *then there exists a constant* C *independent of* p, k *such that*

$$
\sum_{i=0}^{p-k} \frac{S_i^2}{(i+1)^2(i+k)(p-k-i+1)(p+k+i)} \leq C \sum_{i=0}^{p-k} \frac{v_i^2}{(i+k)(p-k-i+1)(p+k+i)}
$$

Proof. Since the proof technique is the same as Lemma 4.10, we will only tersely discuss the details below.

As before, split the inequality into two, similar to (4.24) and (4.25). For the first sum, we set $a_i = |v_i|$, $b_i^{-1} = (i+1)^2(i+k)(p-k-i+1)(p+k+i)$ and $c_i^{-1} = (i+k)(p-k-i+1)(p+k+i)$ for $i = 0, \dots, \frac{p-k}{2}$. Then, $\sum_{i=0}^{n} c_i^{-1} \leq (p+k)(p-k) \sum_{i=0}^{n} (i+k) \approx (p+k)(p-k)(n^2+kn)$ and the following calculation gives that A is bounded:

$$
\begin{aligned}
A^2 &\approx \sup_{n\in[0,\frac{p-k}{2}]} (p+k)(p-k)(n^2+kn) \sum_{i=n}^{\frac{p-k}{2}} \frac{1}{(i+1)^2(i+k)(p-k-i+1)(p+k+i)} \\
&\leq \sup_{n\in[0,\frac{p-k}{2}]} (n^2+kn) \int_n^{(p-k)/2} \frac{1}{(x+1)^2(x+k)}\, dx \\
&\leq \sup_{n\in[0,\frac{p-k}{2}]} n^2 \int_n^{\frac{p-k}{2}} \frac{1}{(x+1)^3}\, dx + kn \int_n^{\frac{p-k}{2}} \frac{1}{(x+1)^2(x+k)}\, dx < \infty.
\end{aligned}
$$

For the second sum, first transform the sum to start the index 0 again. Next, set $a_i = |v_{p-k-i}|$, $b_i^{-1} = (p-k-i+1)^2(p-i)(2p-i)(i+1)$, $c_i^{-1} = (p-i)(2p-i)(i+1)$ for $i = 0, \dots, \frac{p-k}{2} - 1$. Calculating, we have $\sum_{i=0}^{n} c_i^{-1} \leq p^2 \sum_{i=0}^{n} (i+1) \approx p^2 n^2$ and thus

$$
\begin{aligned}
A^2 &\approx \sup_{n\in[0,\frac{p-k}{2}-1]} p^2n^2 \sum_{i=n}^{\frac{p-k}{2}-1} \frac{1}{(p-k-i+1)^2(p-i)(2p-i)(i+1)} \\
&\leq \sup_{n\in[0,\frac{p-k}{2}-1]} pn^2 \int_n^{\frac{p-k}{2}} \frac{1}{(p-k-(p-k)/2+1)^2(p-(p-k)/2)(x+1)}\,dx \\
&\approx \sup_{n\in[0,\frac{p-k}{2}-1]} \frac{pn^2}{(p-k)^2(p+k)} \log\left(\frac{p-k}{2n}\right) < \infty. \qquad \square
\end{aligned}
$$

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55 of National Bureau of Standards Applied Mathematics Series, For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] M. Ainsworth and J. Coyle, *Conditioning of hierarchic p-version Nédélec elements on meshes of curvilinear quadrilaterals and hexahedra*, SIAM J. Numer. Anal., 41 (2003), pp. 731–750.

[3] ———, *Hierarchic finite element bases on unstructured tetrahedral meshes*, International journal for numerical methods in engineering, 58 (2003), pp. 2103–2130.

[4] M. Ainsworth and S. Jiang, *Preconditioning the mass matrix for high order finite element approximation on triangles*, SIAM J. Numer. Anal., 57 (2019), pp. 355–377.

[5] M. Ainsworth, S. Jiang, and M. A. Sanchéz, *An $\mathcal{O}(p^3)$ hp-version fem in two dimensions: Preconditioning and post-processing*, Computer Methods in Applied Mechanics and Engineering, 350 (2019), pp. 766–802.

[6] T. Apel, *Anisotropic finite elements: local estimates and applications*, Advances in Numerical Mathematics, B. G. Teubner, Stuttgart, 1999.

[7] B. Beckermann, S. A. Goreinov, and E. E. Tyrtyshnikov, *Some remarks on the elman estimate for gmres*, SIAM journal on Matrix Analysis and Applications, 27 (2005), pp. 772–778.

[8] T. F. Chan and T. P. Mathew, *Domain decomposition algorithms*, in Acta numerica, 1994, Acta Numer., Cambridge Univ. Press, Cambridge, 1994, pp. 61–143.

[9] L. Demkowicz, *Computing with hp-adaptive finite elements. Vol. 1*, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2007. One and two dimensional elliptic and Maxwell problems, With 1 CD-ROM (UNIX).

[10] Z. Ditzian, *Multivariate Bernstein and Markov inequalities*, J. Approx. Theory, 70 (1992), pp. 273–283.

[11] H. C. Elman, *Iterative methods for large, sparse, nonsymmetric systems of linear equations*, PhD thesis, Yale University New Haven, Conn, 1982.

[12] R. Freund and S. Ruscheweyh, *On a class of chebyshev approximation problems which arise in connection with a conjugate gradient type method*, Numerische Mathematik, 48 (1986), pp. 525–542.

[13] P. Gray and S. Scott, *Autocatalytic reactions in the isothermal, continuous stirred tank reactor: Oscillations and instabilities in the system a + 2b - 3b; b - c*, Chemical Engineering Science, 39 (1984), pp. 1087 – 1097.

[14] A. E. Green and W. Zerna, *Theoretical elasticity*, Dover Publications, Inc., New York, second ed., 1992.

[15] B. Guo and W. Cao, *An additive Schwarz method for the h-p version of the finite element method in three dimensions*, SIAM J. Numer. Anal., 35 (1998), pp. 632–654.

[16] N. Hu, X.-Z. Guo, and I. Katz, *Bounds for eigenvalues and condition numbers in the p-version of the finite element method*, Mathematics of Computation of the American Mathematical Society, 67 (1998), pp. 1423–1450.

[17] R. Jiwari and J. Yuan, *A computational modeling of two dimensional reaction–diffusion brusselator system arising in chemical processes*, Journal of Mathematical Chemistry, 52 (2014), pp. 1535–1551.

[18] G. E. Karniadakis and S. J. Sherwin, *Spectral/hp element methods for computational fluid dynamics*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, second ed., 2005.

[19] A. Kufner, L. Maligranda, and L.-E. Persson, *The Hardy inequality*, Vydavatelský Servis, Plzeň, 2007. About its history and some related results.

[20] T. Lyche and K. Scherer, *On the p-norm condition number of the multivariate triangular Bernstein basis*, J. Comput. Appl. Math., 119 (2000), pp. 259–273. Dedicated to Professor Larry L. Schumaker on the occasion of his 60th birthday.

[21] J.-F. Maitre and O. Pourquier, *Conditionnements et préconditionnements diagonaux pour la p-version des méthodes d'éléments finis pour des problèmes elliptiques du second ordre*, C. R. Acad. Sci. Paris Sér. I Math., 318 (1994), pp. 583–586.

[22] J. Mandel, *Two-level domain decomposition preconditioning for the p-version finite element method in three dimensions*, Internat. J. Numer. Methods Engrg., 29 (1990), pp. 1095–1108.

[23] J. M. Melenk, *hp-finite element methods for singular perturbations*, vol. 1796 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2002.

[24] E. T. Olsen and J. Douglas, Jr., *Bounds on spectral condition numbers of matrices arising in the p-version of the finite element method*, Numer. Math., 69 (1995), pp. 333–352.

[25] R. B. W. T. Ozisik, Sevtap, *On the constants in inverse inequalities in L^2*, Technical Report CAAM TR10-19, Rice University, (2010).

[26] L. F. Pavarino and O. B. Widlund, *A polylogarithmic bound for an iterative substructuring method for spectral elements in three dimensions*, SIAM J. Numer. Anal., 33 (1996), pp. 1303–1335.

[27] J. E. Pearson, *Complex patterns in a simple system*, Science, 261 (1993), pp. 189–192.

[28] S. J. Ruuth, *Implicit-explicit methods for reaction-diffusion problems in pattern formation*, J. Math. Biol., 34 (1995), pp. 148–176.

[29] Y. Saad, *Iterative methods for sparse linear systems*, SIAM, 2003.

[30] J. Schöberl, J. M. Melenk, C. Pechstein, and S. Zaglmayr, *Additive Schwarz preconditioning for p-version triangular and tetrahedral finite elements*, IMA J. Numer. Anal., 28 (2008), pp. 1–24.

[31] C. Schwab, *p- and hp-finite element methods*, Numerical Mathematics and Scientific Computation, The Clarendon Press, Oxford University Press, New York, 1998. Theory and applications in solid and fluid mechanics.

[32] C. Schwab and M. Suri, *The p and hp versions of the finite element method for problems with boundary layers*, Math. Comp., 65 (1996), pp. 1403–1429.

[33] B. Szabó and I. Babuška, *Finite element analysis*, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1991.

[34] A. Toselli and X. Vasseur, *Domain decomposition preconditioners of Neumann-Neumann type for hp-approximations on boundary layer meshes in three dimensions*, IMA J. Numer. Anal., 24 (2004), pp. 123–156.

[35] ———, *A numerical study on Neumann-Neumann methods for hp approximations on geometrically refined boundary layer meshes. II. Three-dimensional problems*, M2AN Math. Model. Numer. Anal., 40 (2006), pp. 99–122.

[36] A. Toselli and O. Widlund, *Domain decomposition methods—algorithms and theory*, vol. 34 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2005.

[37] E. Twizell, A. Gumel, and Q. Cao, *A second-order scheme for the "brusselator" reaction–diffusion system*, Journal of Mathematical Chemistry, 26 (1999), pp. 297–316.

[38] T. Warburton and J. S. Hesthaven, *On the constants in hp-finite element trace inverse inequalities*, Comput. Methods Appl. Mech. Engrg., 192 (2003), pp. 2765–2773.

[39] J. Zhu, Y.-T. Zhang, S. A. Newman, and M. Alber, *Application of discontinuous galerkin methods for reaction-diffusion systems in developmental biology*, Journal of Scientific Computing, 40 (2009), pp. 391–418.