SIAM J. NUMER. ANAL.
Vol. 57, No. 1, pp. 355–377

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

MARK AINSWORTH† AND SHUAI JIANG†

Abstract. The problem of preconditioning the p-version mass matrix on meshes of (possibly curvilinear) triangular elements in two dimensions is considered. Through a judicious choice of hierarchical basis, it is shown that a preconditioner involving only diagonal solves on the vertices, edges, and element interiors gives rise to a preconditioned system for which the condition number is bounded independently of the polynomial order p and the mesh size h. The analysis is performed in the framework of an additive Schwarz method and requires the construction of new polynomial extension theorems, similar to those that are used in the analysis of the stiffness matrix. However, in the case of the mass matrix it is necessary to look at traces and extensions from the space L_2 (rather than H^1) and to make sense of the traces of polynomials regarded as functions in L_2. Numerical examples are presented illustrating the performance of the algorithm.

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

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

DOI. 10.1137/18M1182450

1. Introduction. High order finite element methods have been shown, both in theory and in practice [13, 22, 27], to deliver exponential rates of convergence for large classes of problems, including cases where the solutions exhibit boundary layers and singularities [7, 22]. The choice of basis function to be used in the implementation has proved rather problematic from the outset when it was quickly realized that the natural, Lagrange or Peano polynomial, basis gave rise to exponential growth of the condition number [27]. This led to the use of hierarchic bases which, although considerably better conditioned than the Peano basis, still gave condition numbers that grow algebraically with the polynomial order p [3, 19, 20], e.g., as $\mathcal{O}(p^{4d})$ in d-spatial dimensions.

While a judicious choice of basis can help ameliorate ill-conditioning, the construction of an efficient preconditioner offers much better prospects. The domain decomposition preconditioner developed by Babuska et al. [6] was shown to reduce the growth of the condition number of the stiffness matrix to $\mathcal{O}(1 + \log^2 p)$ in two dimensions. Subsequent work extended these ideas to include preconditioners for the stiffness matrix in higher dimensions, hp-version finite element methods, and boundary element methods, along with the use of more efficient approximate solvers on the subspaces [2, 4, 10, 11, 21]. Despite the rather extensive work on the analysis and construction of preconditioners for the *stiffness matrix*, virtually no attention has been paid to the question of preconditioning the *mass matrix*.

One might reasonably ask if there really is an issue given that the mass matrix for the standard h-version finite element method is well known to be uniformly bounded independent of the mesh size h. Nevertheless, just as for the stiffness matrix, the

*Received by the editors April 23, 2018; accepted for publication (in revised form) November 12, 2018; published electronically February 19, 2019.
http://www.siam.org/journals/sinum/57-1/M118245.html
Funding: This work was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) program.
†Division of Applied Mathematics, Brown University, Providence, RI 02912 (Mark_Ainsworth@brown.edu, Shuai_Jiang@brown.edu).

condition number for the mass matrix for the p-version finite element method is known to grow algebraically with the polynomial order [3, 19, 20].

The need to solve linear systems involving the mass matrix is easy to underestimate. Explicit (and also implicit) time discretization schemes immediately spring to mind and require the inversion of the mass matrix at each time step. However, the need to efficiently invert the mass matrix also arises in less obvious situations including the construction of preconditioners for mixed finite element discretization of the Stokes equations [24]. The linear systems that arise from singularly perturbed problems and plate models for thin elastic bodies have the structure of a mass matrix plus a small multiple of the stiffness matrix meaning, to a large extent, that the system essentially behaves like a mass matrix. It is easy to forget that the mass matrix (or a lumped version) is routinely used as a smoothener for multigrid solvers [8] for the h-version, without causing any eyebrows to be raised, thanks to the fact that the mass matrix is uniformly bounded for the h-version.

The construction of efficient, domain decomposition type preconditioners for the p-version mass matrix is of practical interest, particularly when one turns to applications beyond Poisson type problems, and this has not escaped the attention of the community completely. Early (unpublished) work of Smith [26] looked at preconditioners for the p-version mass matrix quadrilateral elements in two dimensions using tensor product type arguments.

The present work considers the problem of preconditioning the p-version mass matrix on meshes of (possibly curvilinear) triangular elements in two dimensions. Through a judicious choice of hierarchical basis, *it is shown that a preconditioner involving only diagonal solves on the vertices, edges, and element interiors gives rise to a preconditioned system for which the condition number is bounded independently of the polynomial order p and the mesh size h.* The analysis is performed in the framework of an additive Schwarz method (ASM) and requires the construction of new polynomial extension theorems, similar to those that were derived in the analysis of the stiffness matrix in [6]. However, in the case of the mass matrix it is necessary to look at traces and extensions from the space L_2 (rather than H^1) and to make sense of the traces of polynomials regarded as functions in L_2.

The remainder of the paper is organized as follows. In section 2, we define the basis functions on a simplex. In section 3, we present the preconditioner, analyze its cost, and state the main theorem. In section 4, we present several illustrative numerical examples. In section 5, we use domain decomposition techniques to prove the key theorems. In section 6, we prove the technical lemmas and estimates required. We finish in section 7 with a conclusion.

2. Basis functions.

2.1. Basis functions on a triangle.

Let T be the reference triangle in $\mathbb{R}^2$ with vertices $v_1 = (-1,-1), v_2 = (1,-1), v_3 = (-1,1)$, and let the edges of T be denoted by γ_i for $i = 1,2,3$ such that γ_i is opposite of vertex v_i; see Figure 1. Let $p \geq 3$ be a given integer which is fixed throughout, and let $\mathbb{P}_p(T) = \operatorname{span}\{x^\alpha y^\beta : 0 \leq \alpha, \beta, \alpha+\beta \leq p\}$ denote the space of polynomials of total degree p on T. Finally, for $i = 1,2,3$ we let $\lambda_i \in \mathbb{P}_1(T)$ be the barycentric coordinates on T, i.e., the unique polynomial such that $\lambda_i(v_j) = \delta_{ij}$.

The classical Jacobi polynomials on $[-1,1]$ are denoted by $P_n^{(\alpha,\beta)}$, where n is the order of the polynomial and $\alpha, \beta > -1$ are weights [1]. These will be used to define the basis functions on triangle T as follows.

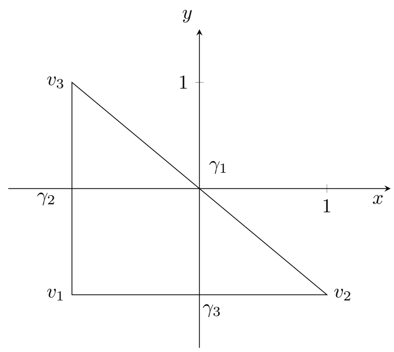

FIG. 1. *Figure of reference triangle T.*

Interior basis functions. The orthogonalized, interior modified principal functions [17] are given by

$$\psi_{ij}(x,y) = \frac{1-s}{2}\frac{1+s}{2}P_{i-1}^{(2,2)}(s)\left(\frac{1-t}{2}\right)^{i+1}\frac{1+t}{2}P_{j-1}^{(2i+3,2)}(t)$$

for $1 \le i, j, i+j \le p-1$, where

$$s = \frac{\lambda_2 - \lambda_1}{1-\lambda_3}, \quad t = 2\lambda_3 - 1$$

and $\lambda_1, \lambda_2, \lambda_3$ are the barycentric coordinates of $(x,y) \in T$. Note that $\{\psi_{ij}\}$ vanishes on the boundary of T and gives a basis for $\mathbb{P}_p(T) \cap H_0^1(T)$.

Edge basis functions. On edge γ_1, we define

$$\chi_n^{(1)}(x,y) = 4\lambda_2\lambda_3 P_n^{(2,2)}(\lambda_3 - \lambda_2)$$

for $n = 0, \ldots, p-2$ with $(x,y) \in T$. We note that the factor $\lambda_2\lambda_3$ means that $\chi_n^{(1)}$ vanishes on edges γ_2 and γ_3. The basis functions $\chi_n^{(2)}, \chi_n^{(3)}$ on edges γ_2, γ_3 are defined in an analogous fashion. The key property dictating this particular choice of basis is that $\chi_n^{(i)}|_{\gamma_i} = (1-s^2)P_n^{(2,2)}(s)$, where $s \in [-1,1]$ is a parametrization of γ_i.

Vertex basis functions. On vertex v_i for $i = 1, 2, 3$, we define

$$\varphi_i(x,y) = \frac{(-1)^{\lfloor p/2 \rfloor + 1}}{\lfloor p/2 \rfloor}\lambda_i P_{\lfloor p/2 \rfloor - 1}^{(1,1)}(1 - 2\lambda_i), \qquad (x,y) \in T.$$

Note that $\varphi_i(v_j) = \delta_{ij}$. One could replace $\lfloor p/2 \rfloor$ by p and still obtain a basis for $\mathbb{P}_p(T)$. The reason for choosing $\lfloor p/2 \rfloor$ rather than simply p will become clear later.

It is not difficult to verify that the functions defined above are linearly independent. Moreover, there are 3 degrees of freedom (dofs) from the vertices, $3p-3$ dofs from the edges, and $\frac{1}{2}\left(p^2 - 3p + 2\right)$ from the interior of T, which sums to $\frac{1}{2}(p+1)(p+2) = \dim \mathbb{P}_p(T)$. Hence, we have a basis for $\mathbb{P}_p(T)$ with the following decomposition:

$$\mathbb{P}_p(T) = \operatorname{span}\{\varphi_i\}_{i=1}^3 \oplus \bigoplus_{i=1}^3 \operatorname{span}\{\chi_n^{(i)}\}_{n=0}^{p-2} \oplus \operatorname{span}\{\psi_{ij}\}_{1\le i,j,i+j\le p-1}. \tag{2.1}$$

This basis bears some similarities to existing [23] bases used for high order FEM and differs slightly in the choice of edge functions, but it uses quite nonstandard vertex functions. Our choice of vertex function is crucial in what follows. For instance, if the usual hat functions were to be used for the vertices, then the resulting preconditioner would result in a condition number which grows as $\mathcal{O}(p^2)$ (see Figure 2).

We enumerate the basis functions in the following order:

1. the vertex functions $\{\varphi_i\}_{i=1}^3$,
2. the edge functions $\{\chi_n^{(1)}\}_{n=0}^{p-2}$, $\{\chi_n^{(2)}\}_{n=0}^{p-2}$, $\{\chi_n^{(3)}\}_{n=0}^{p-2}$,
3. the remaining dofs correspond to $\{\psi_{ij}\}_{1\le i,j,i+j\le p-1}$;

then the mass matrix on T will have a block form

$$\hat{\mathbf{M}} = \begin{bmatrix} \hat{\mathbf{M}}_{VV} & \hat{\mathbf{M}}_{VE} & \hat{\mathbf{M}}_{VI} \\ \hat{\mathbf{M}}_{EV} & \hat{\mathbf{M}}_{EE} & \hat{\mathbf{M}}_{EI} \\ \hat{\mathbf{M}}_{IV} & \hat{\mathbf{M}}_{IE} & \hat{\mathbf{M}}_{II} \end{bmatrix},$$

where $\hat{\mathbf{M}}_{VV} = [\int_T \varphi_i\varphi_j\,dx]$ for $i,j = 1,2,3$ and $\hat{\mathbf{M}}_{VE} = [\int_T \varphi_i\chi_n^{(j)}\,dx]$ for $i,j = 1,2,3$ and $n = 0,\dots,p-2$, etc. Likewise, the element load vector $\vec{f}$ and solution vector $\vec{x}$ take the partitioned forms

$$\vec{f} = \begin{bmatrix} \vec{f}_V \\ \vec{f}_E \\ \vec{f}_I \end{bmatrix} \text{ and } \vec{x} = \begin{bmatrix} \vec{x}_V \\ \vec{x}_E \\ \vec{x}_I \end{bmatrix}.$$

The condition number of the mass matrix $\hat{\mathbf{M}}$ grows as $\mathcal{O}(p^4)$; Figure 2 shows the variation of the condition number versus p. If diagonal scaling is applied as a preconditioner for $\hat{\mathbf{M}}$, then the condition number now grows as $\mathcal{O}(p^2)$; Figure 2 also shows the condition number of the diagonally scaled mass matrix (denoted as $\hat{\mathbf{M}}_S$). Our objective is to construct a preconditioner for which the condition number remains bounded.

2.2. Basis functions on partitions. Let Ω be a bounded two-dimensional domain, and let $\mathcal{T}$ be a triangulation of Ω. We assume that each element $K \in \mathcal{T}$ is the image of the reference element T under a bijective map $\mathcal{F}_K$ (not necessarily linear) such that the Jacobian $D\mathcal{F}_K$ is bounded uniformly in the sense that there exist nonnegative constants θ, Θ such that for all $K \in \mathcal{T}$ there holds

$$\theta|K| \le |D\mathcal{F}_K| \le \Theta|K|. \tag{2.2}$$

We remark that this condition places no constraints on the shape regularity of the mesh and, in particular, allows for "needle" elements.

The basis functions on each element $K \in \mathcal{T}$ are defined in terms of the basis functions on the reference element in the usual way; for example, the first vertex basis functions is defined as

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

Thanks to the decomposition of the basis into interior contributions and boundary contributions that are only supported on a single entity (i.e., edge or vertex), C^0 global conformity is enforced by matching the corresponding edge and vertex functions.

3. Preconditioner and statement of main theorem.

3.1. Preconditioning on the reference element. We begin by constructing a preconditioner for the mass matrix $\hat{\mathbf{M}}$ on the reference element T. Let $\mathbf{I}_3$ be the 3×3 identity matrix, $\hat{\mathbf{D}}_{VV} = \frac{1}{p^4}\mathbf{I}_3$ and

$$\hat{\mathbf{D}}_{EE} = \text{block diag}(\hat{\mathbf{D}}_{EE}^{(1)}, \hat{\mathbf{D}}_{EE}^{(2)}, \hat{\mathbf{D}}_{EE}^{(3)}),$$

where $\hat{\mathbf{D}}_{EE}^{(i)}, i = 1, 2, 3$ is the diagonal matrix $\hat{\mathbf{D}}_{EE}^{(i)} = \text{diag}(q_j)$, with

$$\begin{aligned} q_j &:= \frac{2}{(p+4+j)(p-j+1)} \int_{-1}^{1} (1-x^2)^2 P_j^{(2,2)}(x)^2 \, dx \\ &= \frac{64(j+1)(j+2)}{(p+4+j)(p-j+1)(2j+5)(j+3)(j+4)} \end{aligned} \tag{3.1}$$

for $j = 0, \ldots, p-2$. We define our preconditioner, in the case of the reference element, in terms of its action when applied to a vector $\vec{f}$ in Algorithm 3.1.

Algorithm 3.1 Preconditioner on the reference element.

Require: $\hat{\mathbf{M}}, \vec{f}$ as partitioned in section 2

1: **function**
2: $\quad \vec{x}_I := \hat{\mathbf{M}}_{II}^{-1} \vec{f}_I$ ▷ Interior solve
3: $\quad \vec{x}_E := \hat{\mathbf{D}}_{EE}^{-1} \left(\vec{f}_E - \hat{\mathbf{M}}_{EI} \vec{x}_I \right)$ ▷ Edges solve
4: $\quad \vec{x}_V := \hat{\mathbf{D}}_{VV}^{-1} \left(\vec{f}_V - \hat{\mathbf{M}}_{VI} \vec{x}_I \right)$ ▷ Vertices solve
5: $\quad \vec{x}_I := \vec{x}_I - \hat{\mathbf{M}}_{II}^{-1} \hat{\mathbf{M}}_{IV} \vec{x}_V - \hat{\mathbf{M}}_{II}^{-1} \hat{\mathbf{M}}_{IE} \vec{x}_E$ ▷ Interior correction
6: $\quad$ **return** $\vec{x} := \vec{x}_I + \vec{x}_E + \vec{x}_V$
7: **end function**

Direct manipulation reveals that Algorithm 3.1 defines a linear mapping $\vec{f} \to \vec{x} := \hat{\mathbf{P}}^{-1}\vec{f}$, where $\hat{\mathbf{P}}^{-1} = \hat{\mathbf{Q}}^{-T}\mathbf{D}^{-1}\hat{\mathbf{Q}}^{-1}$,

$$\hat{\mathbf{Q}} := \begin{bmatrix} \mathbf{I} & 0 & \hat{\mathbf{M}}_{VI}\hat{\mathbf{M}}_{II}^{-1} \\ 0 & \mathbf{I} & \hat{\mathbf{M}}_{EI}\hat{\mathbf{M}}_{II}^{-1} \\ 0 & 0 & \mathbf{I} \end{bmatrix} \text{ and } \mathbf{D} := \begin{bmatrix} \hat{\mathbf{D}}_{VV} & 0 & 0 \\ 0 & \hat{\mathbf{D}}_{EE} & 0 \\ 0 & 0 & \hat{\mathbf{M}}_{II} \end{bmatrix}.$$

Clearly, $\hat{\mathbf{Q}}$ and $\mathbf{D}$ are invertible, hence

$$\hat{\mathbf{P}} = \hat{\mathbf{Q}}\mathbf{D}\hat{\mathbf{Q}}^T. \tag{3.2}$$

We now state a key result.

THEOREM 3.1. *There exist positive constants $\hat{c}$ and $\hat{C}$ independent of p such that $\hat{c}\hat{\mathbf{P}} \leq \hat{\mathbf{M}} \leq \hat{C}\hat{\mathbf{P}}$.*[1] *Hence,*

$$\text{cond}(\hat{\mathbf{P}}^{-1}\hat{\mathbf{M}}) \leq \frac{\hat{C}}{\hat{c}}.$$

The proof of Theorem 3.1 is postponed to section 5.

[1]We use the notation that $\mathbf{A} \leq \mathbf{B}$ implies $\mathbf{B} - \mathbf{A}$ is semipositive definite.

3.2. Preconditioning on a mesh. The global mass matrix $\mathbf{M}$ on a partition $\mathcal{T}$ is obtained by the standard finite element subassembly procedure

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

where $\mathbf{M}_K$ is the element mass matrix and $\mathbf{\Lambda}_K$ the local assembly matrix. For the global mass matrix, we assume the dofs are numbered in a fashion similar to the one used on a single element:

1. vertex basis dofs are first in any order,
2. edge basis dofs grouped by the edge they are supported on, and ordered by the index on the Jacobi polynomial,
3. interior basis dofs grouped by the element on which they are supported.

Thanks to (2.2), it follows that

$$c\frac{|K|}{|T|}\hat{\mathbf{M}} \leq \mathbf{M}_K \leq C\frac{|K|}{|T|}\hat{\mathbf{M}} \qquad \forall K \in \mathcal{T},$$

where the constants c and C depend only on θ and Θ. By the same token, we define a local preconditioner on K in terms of $\hat{\mathbf{P}}$,

$$\mathbf{P}_K = \frac{|K|}{|T|}\hat{\mathbf{P}} = \frac{|K|}{|T|}\hat{\mathbf{Q}}\mathbf{D}\hat{\mathbf{Q}}^T, \tag{3.3}$$

where the second equality follows from (3.2). The global preconditioner $\mathbf{P}$ is then obtained using subassembly to give

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

Let the local assembly matrix $\mathbf{\Lambda}_K$ be written in block form

$$\mathbf{\Lambda}_K = \begin{bmatrix} \mathbf{\Lambda}_{K,V} \\ \mathbf{\Lambda}_{K,E} \\ \mathbf{\Lambda}_{K,I} \end{bmatrix},$$

where the blocks correspond to the vertex, edge, and interior basis functions on element K, and let

$$\mathbf{Q} = \begin{bmatrix} \mathbf{I} & 0 & \mathring{\mathbf{M}}_{VI}(\mathring{\mathbf{M}}_{II})^{-1} \\ 0 & \mathbf{I} & \mathring{\mathbf{M}}_{EI}(\mathring{\mathbf{M}}_{II})^{-1} \\ 0 & 0 & \mathbf{I} \end{bmatrix},$$

where $\mathring{\mathbf{M}}_{EI} = \sum_{K \in \mathcal{T}} \mathbf{\Lambda}_{K,E} \hat{\mathbf{M}}_{EI} \mathbf{\Lambda}_{K,I}^T$ with $\mathring{\mathbf{M}}_{II}, \mathring{\mathbf{M}}_{VI}$ defined analogously. Observe that if the physical elements K are all affine images of the reference element, then $\mathring{\mathbf{M}}_{II}, \mathring{\mathbf{M}}_{EI}$ will coincide with the global mass matrix blocks $\mathbf{M}_{II}, \mathbf{M}_{EI}$.

The following identity will prove useful in deducing the action of $\mathbf{P}^{-1}$.

LEMMA 3.2. *For any element $K \in \mathcal{T}$, we have that*

$$\mathbf{\Lambda}_K \hat{\mathbf{Q}} = \mathbf{Q}\mathbf{\Lambda}_K. \tag{3.4}$$

Proof. It is clear that $\mathbf{\Lambda}_K \hat{\mathbf{Q}} \vec{f} = \mathbf{Q}\mathbf{\Lambda}_K \vec{f}$ if $\vec{f} = [\vec{f}_V; \vec{f}_E; \vec{0}]^2$ since, in that case,

$$\mathbf{\Lambda}_K \hat{\mathbf{Q}}[\vec{f}_V; \vec{f}_E; \vec{0}] = [\mathbf{\Lambda}_{K,V} \vec{f}_V; \mathbf{\Lambda}_{K,E} \vec{f}_E; \vec{0}] = \mathbf{Q}\mathbf{\Lambda}_K [\vec{f}_V; \vec{f}_E; \vec{0}].$$

It remains to show the relation holds for vectors of the form $[\vec{0}; \vec{0}; \vec{f}_I]$. Observe that the interior basis functions are supported on one and only one element. Hence $\mathring{\mathbf{M}}_{II}^{-1} = \sum_{K \in \mathcal{T}} \mathbf{\Lambda}_{K,I} \hat{\mathbf{M}}_{II}^{-1} \mathbf{\Lambda}_{K,I}^T$, and $\mathbf{\Lambda}_{K,I}^T \mathbf{\Lambda}_{K',I} = \delta_{KK'}\mathbf{I}$ for $K, K' \in \mathcal{T}$. Direct computation then shows

$$\mathbf{Q}\mathbf{\Lambda}_K \begin{bmatrix} 0 \\ 0 \\ \vec{f}_I \end{bmatrix} = \begin{bmatrix} \mathring{\mathbf{M}}_{VI} \mathbf{\Lambda}_{K,I} \hat{\mathbf{M}}_{II}^{-1} \vec{f}_I \\ \mathring{\mathbf{M}}_{EI} \mathbf{\Lambda}_{K,I} \hat{\mathbf{M}}_{II}^{-1} \vec{f}_I \\ \mathbf{\Lambda}_{K,I} \vec{f}_I \end{bmatrix} = \begin{bmatrix} \mathbf{\Lambda}_{K,V} \hat{\mathbf{M}}_{VI} \hat{\mathbf{M}}_{II}^{-1} \vec{f}_I \\ \mathbf{\Lambda}_{K,E} \hat{\mathbf{M}}_{EI} \hat{\mathbf{M}}_{II}^{-1} \vec{f}_I \\ \mathbf{\Lambda}_{K,I} \vec{f}_I \end{bmatrix} = \mathbf{\Lambda}_K \hat{\mathbf{Q}} \begin{bmatrix} 0 \\ 0 \\ \vec{f}_I \end{bmatrix}. \qquad \square$$

In view of Lemma 3.2 and (3.3), we can rewrite $\mathbf{P}$ in the form

$$\mathbf{P} = \mathbf{Q} \left(\sum_{K \in \mathcal{T}} \mathbf{\Lambda}_K \frac{|K|}{|T|} \mathbf{D} \mathbf{\Lambda}_K^T \right) \mathbf{Q}^T.$$

Moreover, since $\mathbf{D}$ is diagonal, we can rewrite

$$\sum_{K \in \mathcal{T}} \mathbf{\Lambda}_K \frac{|K|}{|T|} \mathbf{D} \mathbf{\Lambda}_K = \text{block diag}(\mathbf{D}_{VV}, \mathbf{D}_{EE}, \mathring{\mathbf{M}}_{II}),$$

where

$$\mathbf{D}_{VV} = \sum_{K \in \mathcal{T}} \frac{|K|}{|T|} \mathbf{\Lambda}_{K,V} \hat{\mathbf{D}}_{VV} \mathbf{\Lambda}_{K,V}^T \text{ and } \mathbf{D}_{EE} = \sum_{K \in \mathcal{T}} \frac{|K|}{|T|} \mathbf{\Lambda}_{K,E} \hat{\mathbf{D}}_{EE} \mathbf{\Lambda}_{K,E}^T.$$

In particular, note that both $\mathbf{D}_{VV}$ and $\mathbf{D}_{EE}$ are diagonal matrices. It follows that $\mathbf{P}$ is invertible, and the action of $\mathbf{P}^{-1}$ on a global right-hand side is given by Algorithm 3.2. A key property of Algorithm 3.2 is that the global preconditioner requires only diagonal solves over the edges, interior, and vertices.

Algorithm 3.2 Preconditioner for global mass matrix.

Require: $\mathbf{M}$ global mass matrix, $\vec{f}$ residual vector

function
- $\vec{x}_I := \mathring{\mathbf{M}}_{II}^{-1} \vec{f}_I$
- $\vec{x}_E := \mathbf{D}_{EE}^{-1} \left(\vec{f}_E - \mathring{\mathbf{M}}_{EI} \vec{x}_I \right)$
- $\vec{x}_V := \mathbf{D}_{VV}^{-1} \left(\vec{f}_V - \mathring{\mathbf{M}}_{VI} \vec{x}_I \right)$
- $\vec{x}_I := \vec{x}_I - \mathring{\mathbf{M}}_{II}^{-1} \mathring{\mathbf{M}}_{IV} \vec{x}_V - \mathring{\mathbf{M}}_{II}^{-1} \mathring{\mathbf{M}}_{IE} \vec{x}_E$
- **return** $\vec{x} := \vec{x}_I + \vec{x}_E + \vec{x}_V$

end function

The next result complements Theorem 3.1 by showing that $\mathbf{P}$ is a uniform preconditioner for the mass matrix on the entire mesh $\mathcal{T}$.

[2] We use the notation where $[\vec{a}; \vec{b}; \vec{c}]$ denotes the column vector $[\vec{a}^T, \vec{b}^T, \vec{c}^T]^T$.

Corollary 3.3. *There exists a constant C independent of h, p such that*

$$\mathrm{cond}(\mathbf{P}^{-1}\mathbf{M}) \le C.$$

Proof. Bounds (2.2) and a change of variables show that $\theta\hat{\mathbf{M}} \le \mathbf{M}_K \le \Theta\hat{\mathbf{M}}$. Then by standard subassembly and Theorem 3.1

$$\hat{c}\theta\mathbf{P} = \hat{c}\theta \sum_{K\in\mathcal{T}} \mathbf{\Lambda}_K \mathbf{P}_K \mathbf{\Lambda}_K^T \le \sum_{K\in\mathcal{T}} \mathbf{\Lambda}_K \mathbf{M}_K \mathbf{\Lambda}_K^T = \mathbf{M} \le \hat{C}\Theta \sum_{K\in\mathcal{T}} \mathbf{\Lambda}_K \mathbf{P}_K \mathbf{\Lambda}_K^T = \hat{C}\Theta\mathbf{P},$$

where $\hat{c}, \hat{C}$ are the constants from Theorem 3.1. Hence $\mathrm{cond}(\mathbf{P}^{-1}\mathbf{M}) \le \frac{\hat{C}\Theta}{\hat{c}\theta}$. ∎

3.3. Cost of applying the preconditioner. Line 2 to line 4 of Algorithm 3.2 all involve inversion of diagonal matrices. Consequently, each interior block can be inverted at a cost of $\frac{1}{2}(p-1)(p-2)$ operations and each edge block at a cost of $p-1$ operations, and the vertex block costs $3|\mathcal{V}|$ operations, where $|\mathcal{V}|$ is the number of vertices in mesh $\mathcal{T}$. The dominant cost of the algorithm lies in the matrix-vector multiplication $\mathbf{M}_{EI}^{\mathrm{pre}}\vec{x}_I$, which costs $\mathcal{O}(p^3)$ operations; hence the overall cost of our algorithm is $\mathcal{O}(p^3)$.

4. Numerical examples. In this section, we present results obtained by applying Algorithm 3.2 to solve linear algebraic systems arising in some representational examples.

4.1. Condition number on reference triangle. We start by illustrating the performance of the preconditioner on the reference element (see Theorem 3.1). In Figure 2, we plot the condition number of $\hat{\mathbf{M}}$, the condition number of the diagonally scaled mass matrix $\hat{\mathbf{M}}_S$ where

$$\hat{\mathbf{M}}_S = \mathrm{diag}(\hat{\mathbf{M}})^{-1/2}\hat{\mathbf{M}}\ \mathrm{diag}(\hat{\mathbf{M}})^{-1/2},$$

and the condition number of the preconditioned mass matrix $\hat{\mathbf{P}}^{-1/2}\hat{\mathbf{M}}\hat{\mathbf{P}}^{-1/2}$.

Figure 2 also shows the results obtained if the vertex functions in the choice of basis are replaced by the "full-order" vertex basis functions

$$\ddot{\varphi}_i(x,y) = \frac{(-1)^{p+1}}{p}\lambda_i P_{p-1}^{(1,1)}(1-2\lambda_i), \qquad (x,y) \in T,$$

to partially illustrate why the choice $\lfloor p/2 \rfloor$ was made. We will call the preconditioned mass matrix constructed using $\ddot{\varphi}_i$ as $\hat{\mathbf{P}}^{-1/2}\ddot{\mathbf{M}}\hat{\mathbf{P}}^{-1/2}$. It is observed that the condition number no longer remains bounded; see Lemma 6.3 for an explanation. Finally, the figure also shows the results obtained if the vertex functions are replaced with the commonly used hat functions

$$\breve{\varphi}_i(x,y) = \lambda_i, \qquad (x,y) \in T.$$

We call the preconditioned mass matrix constructed using $\breve{\varphi}_i$ as $\hat{\mathbf{P}}_L^{-1/2}\breve{\mathbf{M}}\hat{\mathbf{P}}_L^{-1/2}$; the only difference between $\hat{\mathbf{P}}$ and $\hat{\mathbf{P}}_L$ is a more appropriate scaling for $\hat{\mathbf{D}}_{VV}$. Figure 2 shows the growth of the condition number is of order $\mathcal{O}(p^2)$.

We note that the mass matrix $\hat{\mathbf{M}}$ and the scaled mass matrix $\hat{\mathbf{M}}_S$ both exhibit algebraic growth with the order p, which is typically the case for such basis [3], while, by contrast, the preconditioned system $\hat{\mathbf{P}}^{-1/2}\hat{\mathbf{M}}\hat{\mathbf{P}}^{-1/2}$ remains constant with p as predicted by Theorem 3.1 (with an asymptotic value of 24 as $p \to \infty$).

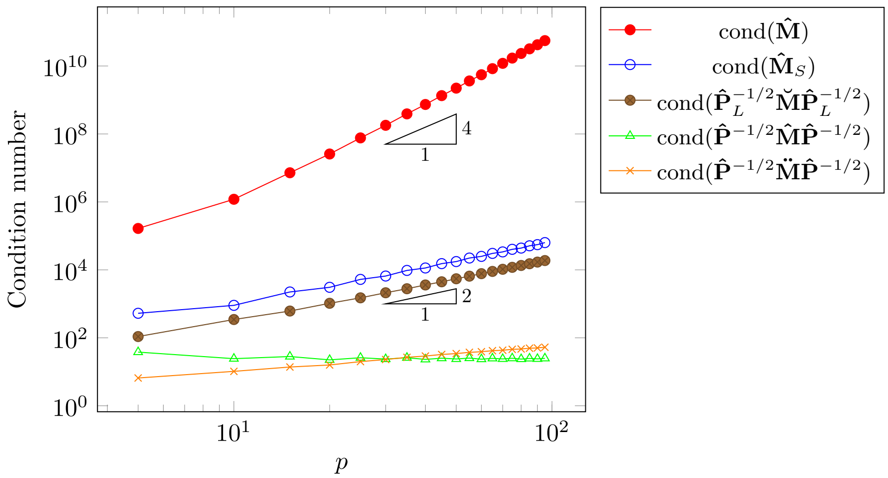

FIG. 2. *The condition numbers of* $\hat{\mathbf{M}}, \hat{\mathbf{M}}_S$*,* $\mathrm{cond}(\hat{\mathbf{P}}_L^{-1/2}\breve{\mathbf{M}}\hat{\mathbf{P}}_L^{-1/2})$*,* $\hat{\mathbf{P}}^{-1/2}\hat{\mathbf{M}}\hat{\mathbf{P}}^{-1/2}$*, and* $\hat{\mathbf{P}}^{-1/2}\ddot{\mathbf{M}}\hat{\mathbf{P}}^{-1/2}$ *are plotted on a log-log axis for* $p = 5, 10, \ldots, 95$*. The algebraic growth of* $\mathrm{cond}(\hat{\mathbf{M}})$ *and* $\mathrm{cond}(\hat{\mathbf{M}}_S)$ *with* p *are consistent with* [3]*, and the boundedness of* $\mathrm{cond}(\hat{\mathbf{P}}^{-1/2}\hat{\mathbf{M}}\hat{\mathbf{P}}^{-1/2})$ *is predicted in Theorem* 3.1*. Finally, we note the importance of our choice of vertex function: the "full-order" vertex basis system* $\mathrm{cond}(\hat{\mathbf{P}}^{-1/2}\ddot{\mathbf{M}}\hat{\mathbf{P}}^{-1/2})$ *grows and the hat functions system* $\mathrm{cond}(\hat{\mathbf{P}}_L^{-1/2}\breve{\mathbf{M}}\hat{\mathbf{P}}_L^{-1/2})$ *exhibits* $\mathcal{O}(p^2)$ *growth.*

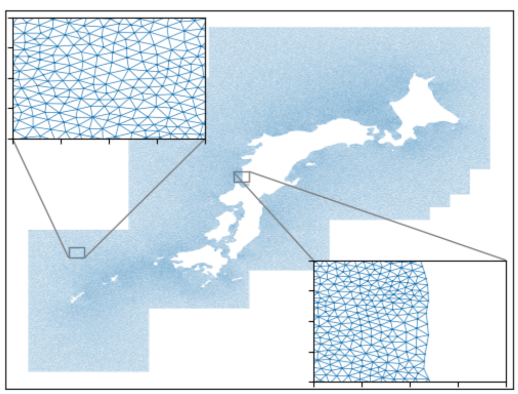

FIG. 3. *Plot of the mesh used to illustrate Corollary* 3.3*; see Table* 1 *for the results.*

4.2. Condition number on multielement mesh. We next illustrate Corollary 3.3 by considering the mesh shown in Figure 3, which consists of 239852 affine elements. We construct the global mass matrix $\mathbf{M}$ explicitly and use ARPACK to approximate the extreme eigenvalues of the preconditioned system to a relative tolerance of 10^{-4}. In Table 1, we display the extreme eigenvalues and condition number of the preconditioned mass matrix on the multielement mesh, along with the corresponding quantities for the preconditioned mass matrix on the reference element. The condition numbers on the multielement mesh are bounded by those on the reference element as predicted by Corollary 3.3 for affine elements.

4.3. Explicit time-stepping. We now illustrate the use of the preconditioner in the numerical solution of the wave equation where the time-stepping scheme requires

Table 1

Table illustrates Corollary 3.3 by comparing the extreme eigenvalues of the global mass matrix $\mathbf{M}$ of the mesh as shown in Figure 3 to the single element case $\hat{\mathbf{M}}$. The eigenvalues are approximated using ARPACK to a relative tolerance of 10^{-4} for $\mathbf{M}$ and to machine precision for $\hat{\mathbf{M}}$.

		Multielement mesh $\mathbf{M}$			Single element $\hat{\mathbf{M}}$		
p	#dof	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\max}/\lambda_{\min}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\max}/\lambda_{\min}$
3	1084371	0.0518	2.6077	50.341	0.0518	2.6124	50.386
4	1925541	0.0922	2.3033	24.982	0.0920	2.3064	25.061
5	3006563	0.0793	2.9154	36.764	0.0791	2.9198	36.887

Table 2

Table illustrates the performance of the preconditioned iterative method of the mass matrix at each time step by displaying the [min, median, max] iteration count of all 3000 PCG solves from using the Nyström method for a period of 10 seconds with a $\Delta t = .01$ on $u_{tt} = \Delta u$ in a uniformly triangulated square. The iteration count does not increase as predicted in Corollary 3.3 and (4.1).

Order	16 elements	64 elements	256 elements
4	[21, 27, 34]	[20, 25, 34]	[17, 23, 31]
8	[17, 23, 29]	[16, 21, 30]	[16, 21, 26]
12	[17, 22, 27]	[16, 18, 26]	[16, 17, 25]
16	[16, 18, 25]	[15, 18, 24]	[15, 15, 23]
20	[16, 18, 24]	[15, 15, 23]	

the inversion of the mass matrix at each step. Let $u(x, y, t)$ be the solution to the wave equation

$$u_{tt} = \Delta u, \qquad (x, y) \in \Omega = [-7, 7] \times [-7, 7], t > 0$$

with Neumann boundary condition; the initial condition [9] is

$$u(x, y, 0) = 4 \tan^{-1} \exp(x + 1 - 2 \operatorname{sech}(y + 7) - 2 \operatorname{sech}(y - 7)), \qquad u_t(x, y, 0) = 0.$$

For the spatial discretization, we use a uniform triangulation of the square. For the time discretization, we use a fourth order Nyström method [15, p. 285], which entails three mass matrix solves per time step; for example, the first substep consists of solving

$$\vec{u}_1^{n+1} := \mathbf{M}^{-1} \left(-\mathbf{S}\vec{u}^n \right),$$

where $\mathbf{S}$ is the stiffness matrix. For each solve, we use the preconditioned conjugate gradient (PCG) with an appropriate initial guess; recall that the error $\vec{e}_k$ at iteration k of PCG satisfies

$$\|\vec{e}_k\| \leq \left(\frac{\sqrt{\kappa - 1}}{\sqrt{\kappa + 1}} \right)^k \|\vec{e}_0\|, \tag{4.1}$$

where κ is the condition number of the preconditioned matrix and $\vec{e}_0$ is the error of the initial iterate [14, p. 636]. In Table 2, we show the minimum, median, and max iteration count of PCG over the entire simulation of 10 seconds with $\Delta t = 0.01$.

Corollary 3.3 and (4.1) guarantee that the iteration count will not increase with p or with h refinement. In fact, we note that the median iteration count actually *decreases* as we increase p and refine h. This is due to (4.1) being an estimate which only relates the condition number to the error bound, but does not take into account the possible improvements from clustering of eigenvalues. Furthermore, the estimate does not take into account a good initial iterate, which improves as we increase the number of dofs.

4.4. Implicit time-stepping. Finally, we illustrate the use of the preconditioner in the solution of the heat equation where the time-stepping scheme requires the inversion of a perturbed mass matrix at each step. Let $u(x, y, t)$ be the solution to the heat equation

$$u_t = \Delta u, \qquad (x, y) \in \Omega = [-1, 1] \times [-1, 1], t > 0$$

with Neumann boundary condition; we use a simple initial condition

$$u(x, y, 0) = \exp(-(x^2 + y^2)).$$

The time-stepping scheme we use is the Crank–Nicolson method:

$$\left(\mathbf{M} + \frac{\Delta t}{2}\mathbf{S}\right) \vec{u}^{n+1} = \left(\mathbf{M} - \frac{\Delta t}{2}\mathbf{S}\right) \vec{u}^{n},$$

where $\mathbf{S}$ is the stiffness matrix. By Schmidt's inequality [16], there exists a c independent of p, h such that

$$0 \leq \mathbf{S} \leq c\frac{p^4}{h^2}\mathbf{M} \implies \mathbf{M} \leq \mathbf{M} + \frac{\Delta t}{2}\mathbf{S} \leq \left(1 + \frac{1}{2}c\Delta t\frac{p^4}{h^2}\right)\mathbf{M}. \tag{4.2}$$

The preconditioned system will have condition number of

$$\operatorname{cond}\left(\mathbf{P}^{-1}\left(\mathbf{M} + \frac{1}{2}\Delta t\mathbf{S}\right)\right) = \mathcal{O}\left(1 + \Delta t\frac{p^4}{h^2}\right). \tag{4.3}$$

Observe that if we were to use a fully explicit scheme, then the CFL condition is $\Delta t \sim \frac{h^2}{p^4}$ thanks again to Schmidt's inequality being sharp. If we use the choice $\Delta t \sim \frac{h^2}{p^4}$ for the implicit scheme, then (4.3) shows that the iteration count will not increase as we increase p. In practice, however, one generally chooses $\Delta t \sim \frac{h^2}{p^2}$, in which case (4.3) shows that the condition number will grow at a rate of at most $\mathcal{O}(p^2)$; hence the iteration count will also increase. These conclusions are illustrated in Table 3. In the first two columns, we start with an initial iterate of $\vec{0}$ in each PCG method. In the other two columns, we use the solution from the previous time step as the initial iterate, which results in drastic decreases in iteration counts.

We remark (4.3) could be improved to $\mathcal{O}((1+\log^2 p)(1+\log^2(p/h)))$ by combining Algorithm 3.2 with a domain decomposition preconditioner for the stiffness matrix [2] but this would require a significant increase in computational cost.

TABLE 3

Table illustrates the performance of the preconditioned iterative method to the matrix resulting from Crank–Nicolson scheme by displaying the [min, median, max] iteration count of all PCG solves from using Crank–Nicolson for a period of 1 *seconds on* 16 *elements for* $u_t = \Delta u$ *in a uniformly triangulated square. For the latter two columns, the initial guess is the previous time step. The behavior as we increase* p *is predicted by* (4.3).

	Initial iterate: $\vec{0}$		Initial iterate: $\vec{u}^n$	
p	$\Delta t \sim \frac{h^2}{p^4}$	$\Delta t \sim \frac{h^2}{p^2}$	$\Delta t \sim \frac{h^2}{p^4}$	$\Delta t \sim \frac{h^2}{p^2}$
4	[35, 36, 37]	[35, 36, 37]	[34, 34, 36]	[34, 34, 36]
8	[38, 39, 39]	[66, 67, 73]	[9, 17, 35]	[49, 51, 73]
12	[34, 35, 35]	[87, 91, 103]	[4, 8, 29]	[51, 55, 101]
16	[32, 33, 33]	[108, 114, 127]	[2, 7, 24]	[48, 55, 124]
20	[16, 19, 19]	[129, 130, 151]	[1, 1, 9]	[47, 55, 149]

5. Additive Schwarz theory. Thanks to Corollary 3.3, the analysis of the preconditioner reduces to bounding the condition number on the reference element as in Theorem 3.1. Consequently, for the remainder of this article we confine our attention to the reference triangle.

Let $X := \mathbb{P}_p(T)$ be equipped with the standard L^2 inner-product denoted by $(\cdot,\cdot)$ with the respective norm denoted by $\|\cdot\|$, and let $X_I := H_0^1(T) \cap \mathbb{P}_p(T)$ be the interior space equipped with the $L^2(T)$ inner-product. The orthogonal complement of the (closed) subspace X_I in X is denoted by $\widetilde{X}_B$, i.e.,

$$X = X_I \oplus \widetilde{X}_B, \qquad X_I \perp \widetilde{X}_B. \tag{5.1}$$

We begin by exploring the structure of the space $\widetilde{X}_B$. Let $\mathbb{P}_p(\partial T)$ denote the space of traces of $\mathbb{P}_p(T)$ on the boundary ∂T of the reference triangle:

$$\mathbb{P}_p(\partial T) = \{u : u = v|_{\partial T} \text{ for some } v \in \mathbb{P}_p(T)\}. \tag{5.2}$$

The next result shows that there is a one-to-one correspondence between $\widetilde{X}_B$ and $\mathbb{P}_p(\partial T)$.

LEMMA 5.1. *For every $u \in \mathbb{P}_p(\partial T)$, there exists a unique $\widetilde{u} \in \widetilde{X}_B$ which satisfies $\widetilde{u} = u$ on ∂T, and $(\widetilde{u}, v) = 0$ for all $v \in X_I$. Furthermore, $\widetilde{u}$ is a minimal L^2 extension of u in the sense that for all $w \in \mathbb{P}_p(T)$ with $w|_{\partial T} = u$ we have $\|\widetilde{u}\| \leq \|w\|$.*

Proof. Let $u \in \mathbb{P}_p(\partial T)$ be given. According to (5.2), u is equal to the trace of a polynomial in $\mathbb{P}_p(T)$, which we again denote by u. We can construct a $\widetilde{u} \in \widetilde{X}_B$ with the claimed properties as follows.

Let

$$u_I \in X_I : (u_I, v_I) = -(u, v_I) \qquad \forall v_I \in X_I.$$

Set $\widetilde{u} = u + u_I$; clearly $\widetilde{u}|_{\partial T} = u$ and $(\widetilde{u}, v_I) = 0$ for all $v_I \in X_I$; this gives existence. For uniqueness, let $\tilde{w} \in \mathbb{P}_p(T) : \widetilde{w}|_{\partial T} = u, (\widetilde{w}, v_I) = 0$ for all $v_I \in X_I$, then

$$(\tilde{u} - \tilde{w}, v_I) = 0 \qquad \forall v_I \in X_I.$$

Hence $\tilde{u} - \tilde{w} = 0$ as $\tilde{u} - \tilde{w} \in X_I$. The minimal L^2 extension property follows from the Pythagorean identity. ∎

We say that $\widetilde{u}$ is the "minimal L^2 extension" or "minimal extension" of $u \in \mathbb{P}_p(\partial T)$. Lemma 5.1 shows that $\widetilde{u}$ is uniquely determined by the boundary values of u and the degree of the space.

We decompose the space $\widetilde{X}_B$ further. Let $\widetilde{\varphi}_i$ and $\widetilde{\chi}_n^{(i)}$ be the minimal extension, constructed as described in Lemma 5.1, of the vertex basis function and the edge basis function defined in section 2, respectively. Let

$$\widetilde{X}_V = \operatorname{span}\{\widetilde{\varphi}_i : i = 1, 2, 3\}$$

and

$$\widetilde{X}_{E_i} = \operatorname{span}\{\widetilde{\chi}_n^{(i)} : n = 0, \ldots, p-2\}, \qquad i = 1, 2, 3.$$

By the construction of the basis functions on the boundary and, thanks to (2.1) and (5.1), we have

$$X = X_I \oplus \widetilde{X}_V \oplus \bigoplus_{i=1}^{3} \widetilde{X}_{E_i}. \tag{5.3}$$

Let $\vec{\varphi} = [\varphi_1; \varphi_2; \varphi_3]$, where φ_i are the vertex basis functions with $\vec{\psi}$ defined similarly for the interior basis functions, and, using the notation of section 2, define

$$\vec{\widetilde{\varphi}} = \vec{\varphi} - \hat{\mathbf{M}}_{VI}\hat{\mathbf{M}}_{II}^{-1}\vec{\psi}. \tag{5.4}$$

Then for $\vec{u} \in \mathbb{R}^3$, we have for all $X_I \ni w = \vec{w}^T\vec{\psi}$,

$$\begin{aligned}(\vec{u}^T\vec{\widetilde{\varphi}}, w) = \left(\vec{u}^T\vec{\widetilde{\varphi}}, \vec{w}^T\vec{\psi}\right) &= \left(\vec{u}^T(\vec{\varphi} - \hat{\mathbf{M}}_{VI}\hat{\mathbf{M}}_{II}^{-1}\vec{\psi}), \vec{w}^T\vec{\psi}\right) \\ &= \vec{u}^T\hat{\mathbf{M}}_{VI}\vec{w} - \vec{u}^T\hat{\mathbf{M}}_{VI}\hat{\mathbf{M}}_{II}^{-1}\hat{\mathbf{M}}_{II}\vec{w} = 0.\end{aligned}$$

Hence $\{\widetilde{\varphi}_1, \widetilde{\varphi}_2, \widetilde{\varphi}_3\} \in \widetilde{X}_B$, and as a consequence forms a basis for $\widetilde{X}_V$ (since $\widetilde{\varphi}_i|_{\partial T} = \varphi_i|_{\partial T}$). A basis for $\widetilde{X}_{E_i}$ with $i = 1, 2, 3$ can be constructed in the same fashion.

Next, we define the bilinear forms on each subspace in the decomposition (5.3):

- Interior space X_I:

$$a_I(u, w) := (u, w), \qquad u, w \in X_I.$$

- Vertex space $\widetilde{X}_V$:

$$a_V(u, w) := \frac{1}{p^4}\sum_{i=1}^{3} u(v_i)w(v_i), \qquad u, w \in \widetilde{X}_V,$$

  where v_1, v_2, v_3 are the vertices of T.
- Edge spaces $\widetilde{X}_{E_i}$ $(i = 1, 2, 3)$:

$$a_{E_i}(u, w) := \sum_{n=0}^{p-2} q_n\mu_n(u)\mu_n(w), \qquad u, w \in \widetilde{X}_{E_i},$$

  with q_n defined as in (3.1), and μ_n is the weighted moment given by

$$\mu_n(u) := \frac{(2n+5)(n+3)(n+4)}{32(n+1)(n+2)}\int_{-1}^{1} \chi_n^{(i)}(x)u(x)\,dx,$$

  where we use a linear parametrization such that $\gamma_i = [-1, 1]$.

The spaces and inner-products defined above give rise to an ASM preconditioner [12, 25, 28] whose action on a given residual $f \in X$ is defined as follows:

(i)	$u_I \in X_I : a_I(u_I, v_I) = (f, v_I)\ \forall v_I \in X_I$.
(ii)	$u_V \in \widetilde{X}_V : a_V(u_V, v_V) = (f, v_V)\ \forall v_V \in \widetilde{X}_V$.
(iii)	For $i = 1, 2, 3$, $u_{E_i} \in \widetilde{X}_{E_i} : a_{E_i}(u_{E_i}, v_{E_i}) = (f, v_{E_i})\ \forall v_{E_i} \in \widetilde{X}_{E_i}$.
(iv)	$u := u_I + u_V + \sum_{i=1}^{3} u_{E_i}$ is our solution.

5.1. Matrix formulation of the ASM. In practice, it is convenient to reformulate steps (i)–(iv) in terms of matrix operations.

(1) Recall that $X_I = \operatorname{span}\{\psi_{ij}\}$ and let $u_I = \vec{u}_I^T\vec{\psi}$, where $\vec{\psi}$ is the column vector of all the interior basis functions. The matrix form of (i) is

$$\hat{\mathbf{M}}_{II}\vec{u}_I = a_I(u_I, \vec{\psi}) = (f, \vec{\psi}) = \vec{f}_I.$$

(2) Let $u_V = \vec{u}_V^T \vec{\widetilde{\varphi}}$, where $\vec{\widetilde{\varphi}}$ is the basis for $\widetilde{X}_V$ in column form. As $\widetilde{\varphi}_i(v_j) = \delta_{ij}$, we have

$$\frac{1}{p^4}\mathbf{I}_{VV}\vec{u}_V = a_V(u_V, \vec{\widetilde{\varphi}}) = (f, \vec{\widetilde{\varphi}}).$$

Inserting identity (5.4) in the right-hand side gives

$$\begin{aligned}(f, \vec{\widetilde{\varphi}}) &= (f, \vec{\varphi}) - \mathbf{M}_{VI}\mathbf{M}_{II}^{-1}(f, \vec{\psi}) \\ &= \vec{f}_V - \hat{\mathbf{M}}_{VI}\hat{\mathbf{M}}_{II}^{-1}\vec{f}_I.\end{aligned}$$

(3) Let $u_{E_1} = \vec{u}_{E_1}^T \vec{\widetilde{\chi}}$, where $\vec{\widetilde{\chi}}$ is the basis for $\widetilde{X}_{E_1}$ in column form. By the orthogonality properties of $P_i^{(2,2)}(x)$ in (3.1), the weighted moments in $a_V(\cdot,\cdot)$ of (iii) simplify to $\mu_n(\widetilde{\chi}_i)\mu_n(\widetilde{\chi}_j) = \delta_{ij}$, and hence we have

$$\hat{\mathbf{D}}_{EE}^{(1)}\vec{u}_{E_1} = a_{E_1}(u_{E_1}, \vec{\widetilde{\chi}}) = (f, \vec{\widetilde{\chi}}).$$

The same reasoning holds for edges γ_2, γ_3. The right-hand-side modification follows from (2).

(4) The vector solution $\vec{x}_V$ to step (ii) corresponds to the function $\tilde{u}_V := \vec{x}_V^T \vec{\widetilde{\varphi}}$. Applying identity (5.4) again, we have

$$\tilde{u}_V = \vec{x}_V^T \left(\vec{\varphi} - \hat{\mathbf{M}}_{VI}\hat{\mathbf{M}}_{II}^{-1}\vec{\psi}\right).$$

Therefore, our minimal energy solution contains interior functions of the form $-\hat{\mathbf{M}}_{II}^{-1}\hat{\mathbf{M}}_{IV}\vec{x}_V$ which we have to add back to $\vec{x}_I$. A similar correction term is needed for the three edge terms.

Theorem 5.2. *The abstract ASM defined above corresponds to Algorithm* 3.1.

Proof. Steps (1), (2), (3), and (4) above correspond to lines 2, 4, 3, and 5, respectively, from Algorithm 3.1. ∎

5.2. Proof of Theorem 3.1. We apply the standard theory [12, 25, 28] for the analysis of ASMs to the scenario as described above. In particular, we will follow the framework as laid out in [28, section 2].

Lemma 5.3 (local stability). *For a constant C independent of p, each of our local bilinear forms is coercive in the sense that*

$$\begin{aligned}(u,u) &= a_I(u,u) && \forall u \in X_I, \\ (u,u) &= a_{E_i}(u,u) && \forall u \in \widetilde{X}_{E_i}, i = 1,2,3, \\ (u,u) &\le 3Ca_V(u,u) && \forall u \in \widetilde{X}_V.\end{aligned}$$

Proof. The first equality holds as X_I is a subspace of X and inherits the inner-product. For $\widetilde{X}_{E_i}$, identity (6.3) of Lemma 6.4 gives us the equality

$$a_{E_i}(u,u) = \sum_{n=0}^{p-2} q_n \mu_n(u)^2 = \|u\|^2.$$

Finally, for $u \in \widetilde{X}_V$, we rewrite $u = \sum_{i=1}^{3} u(v_i)\widetilde{\varphi}_i$. Using the triangle inequality and the estimate $\|\widetilde{\varphi}_i\|^2 \le Cp^{-4}$ of Lemma 6.3, we have

$$\|u\|^2 \le 3\sum_{i=1}^{3}\left\|u(v_i)\widetilde{\varphi}_i\right\|^2 \le \frac{3C}{p^4}\sum_{i=1}^{3}|u(v_i)|^2 = 3Ca_V(u,u). \qquad \square$$

The next result gives an estimate for the largest eigenvalue and is an immediate consequence of the triangle inequality and Lemma 5.3.

Lemma 5.4. *There exists a constant C independent of p such that for all $u \in X$, the unique decomposition*

$$u = u_I + u_V + \sum_{i=1}^{3} u_{E_i},$$

with $u_I \in X_I, u_V \in \widetilde{X}_V, u_{E_i} \in \widetilde{X}_{E_i}$, satisfies

$$\|u\|^2 \le C\left(a_I(u_I,u_I) + a_V(u_V,u_V) + \sum_{i=1}^{3} a_{E_i}(u_{E_i},u_{E_i})\right).$$

The final ingredient is the following bound for the smallest eigenvalue of the additive Schwarz operator, whose proof is the subject of section 6:

Theorem 5.5 (stable decomposition). *For all $u \in X$, with the decomposition as in Lemma 5.4, there exists a constant C independent of p such that*

$$a_I(u_I,u_I) + a_V(u_V,u_V) + \sum_{i=1}^{3} a_{E_i}(u_{E_i},u_{E_i}) \le C\|u\|^2.$$

The proof of Theorem 3.1 is now an immediate consequence of Lemmas 5.3 and 5.4 and Theorem 5.5 thanks to Theorem 2.7 of [28].

6. Technical lemmas. In this section, we present the technical lemmas that were used in the proof of Theorem 3.1. For notational purposes, we let $\|\cdot\|_\omega$ define the L^2 norm over a domain ω, and we shall omit the subscript in the case $\omega = T$, the reference element.

We begin with a bound relating the vertex values of a polynomial to its L^2 norm over the triangle. The constant appearing in Lemma 6.1 is the best one possible; a related result was proved in [29].

Lemma 6.1. *For $u \in \mathbb{P}_p(T)$, we have that*

$$\max_{i\in\{1,2,3\}}|u(v_i)| \le \frac{1}{2\sqrt{2}}(p+1)(p+2)\|u\|.$$

Proof. For $0 \le i,j, i+j \le p$ define

$$\Psi_{ij}(x,y) = \sqrt{\frac{(2i+1)(i+j+1)}{2}}P_i^{(0,0)}(\xi)\left(\frac{1-\eta}{2}\right)^i P_j^{(2i+1,0)}(\eta), \tag{6.1}$$

where $\xi = \frac{2(1+x)}{1-y} - 1$ and $\eta = y$ [17, section 3]. These functions form an orthonormal basis for $\mathbb{P}_p(T)$. Hence, $u \in \mathbb{P}_p(T)$ can be written in the form $u = \sum_{i+j\le p} u_{ij}\Psi_{ij}$ and

$\|u\|^2 = \sum_{i+j\le p} u_{ij}^2$. It suffices to prove the inequality in the case of vertex $(-1,-1)$. Using Cauchy–Schwarz gives

$$|u(-1,-1)|^2 = \left(\sum_{i+j\le p}(-1)^{i+j}u_{ij}\sqrt{\frac{(2i+1)(i+j+1)}{2}}\right)^2$$
$$\le \sum_{i+j\le p} u_{ij}^2 \sum_{i+j\le p}\frac{(2i+1)(i+j+1)}{2} = \frac{1}{8}(p+1)^2(p+2)^2\|u\|^2. \qquad \square$$

Next, we prove an equality needed to bound the minimal extension of the vertex functions.

LEMMA 6.2. *Define*

$$\xi_p(x) = \frac{(-1)^{p+1}}{p(p+1)}P_p'(x)(1-x) = \frac{(-1)^{p+1}}{p}\frac{1-x}{2}P_{p-1}^{(1,1)}(x), \quad x\in[-1,1],$$

where P_p *is the Legendre polynomial. Then*

$$\|\xi_p\|^2_{[-1,1]} = \frac{4}{(p+1)(2p+1)}. \tag{6.2}$$

Proof. We note that $\xi_p(-1)=1, \xi_p(1)=0$, and $\xi_p(x_i)=0$, where $x_i, i=2,\dots,p$, are the roots of $P_p'(x)$. Hence, using the $(p+1)$ point Gauss-Lobatto quadrature gives

$$\int_{-1}^{1}\xi_p^2(x)\,dx = w_1 + \sum_{i=2}^{p} w_i\xi_p^2(x_i) + E,$$

where E is the error term

$$E = -\frac{(p+1)p^3 2^{2p+1}[(p-1)!]^4}{(2p+1)[(2p)!]^3}\frac{\mathrm{d}^{2p}}{\mathrm{d}x^{2p}}\xi_p^2(x)\big|_{x=\eta}, \qquad \eta\in[-1,1],$$

for some $\eta\in[-1,1]$. Direct calculation shows that $E = -\frac{2}{(2p+1)(p+1)p}$ which, along with the fact that $w_1 = \frac{2}{p(p+1)}$, gives the result claimed. $\square$

Using the function defined in Lemma 6.2, we can bound the minimal extensions of the vertex functions.

LEMMA 6.3. *The minimal extension of the vertex basis function of degree* p *satisfies the bound*

$$\frac{c}{p^4} \le \|\widetilde{\varphi}_i\|^2 \le \frac{C}{p^4},$$

where c *and* C *are positive constants independent of* p.

Proof. Without loss of generality, assume that $i=1$, which corresponds to $v_1 = (-1,-1)$ of the reference triangle T. Using the minimal L^2 property of $\widetilde{\varphi}_1$, and $\mathbb{Q}_{\lfloor p/2\rfloor}\subset\mathbb{P}_p$ where $\mathbb{Q}_r = \{x^\alpha y^\beta : 0\le\alpha,\beta\le r\}$, gives

$$\|\widetilde{\varphi}_1\|^2 = \min_{\substack{u=\varphi_1 \text{ on } \partial T\\ u\in\mathbb{P}_p}}\|u\|^2 \le \min_{\substack{u=\varphi_1 \text{ on } \partial T\\ u\in\mathbb{Q}_{\lfloor p/2\rfloor}}}\|u\|^2.$$

Consider the polynomial $\zeta_r \in \mathbb{Q}_{2r}$ defined by

$$\zeta_r(x,y) = \xi_r(x)\xi_r(y) - \xi_r(-x)\xi_r(-y),$$

where $\xi_r(x)$ is defined in Lemma 6.2. By construction, $\zeta_{\lfloor p/2 \rfloor} = \varphi_1$ on ∂T, and using (6.2) gives

$$\left\|\zeta_{\lfloor p/2\rfloor}\right\|^2 = \frac{4(2\lfloor p/2\rfloor - 1)}{\lfloor p/2\rfloor^2(\lfloor p/2\rfloor+1)^2(2\lfloor p/2\rfloor+1)} \le \frac{C}{p^4},$$

which proves the upper bound.

The lower bound is an immediate consequence of Lemma 6.1 (choosing $u = \widetilde{\varphi}_i$). ∎

Remark. The $\lfloor p/2 \rfloor$ order on the vertex functions is crucial here to guarantee that $\mathbb{Q}_{\lfloor p/2\rfloor}$ is a smaller space than $\mathbb{P}_p$. Using p as the order on the Legendre polynomial will result in a growing condition number; see Figure 2.

The next result gives an explicit expression for the norm of a minimal extension of an edge function.

LEMMA 6.4. *Let $u \in \mathbb{P}_p(\gamma)$, a polynomial on edge $\gamma \subset \partial T$ which vanishes at the endpoints, be written in the form*

$$u(x) = (1-x^2)\sum_{i=0}^{p-2} w_i P_i^{(2,2)}(x),$$

where $x \in [-1,1]$ is a parametrization of γ. Then the norm of the minimal energy extension $\widetilde{u} \in \mathbb{P}_p(T)$, satisfying $\widetilde{u} = 0$ on $\partial T \setminus \gamma$ and $u = \widetilde{u}$ on γ, is given by

$$\|\widetilde{u}\|^2 = \sum_{i=0}^{p-2} \frac{2\mu_i w_i^2}{(p+i+4)(p-i-1)}, \tag{6.3}$$

where $\mu_i = \int_{-1}^{1}(1-x^2)^2 P_i^{(2,2)}(x)^2\,dx = \frac{32}{2i+5}\frac{(i+1)(i+2)}{(i+3)(i+4)}$.

Proof. Without loss of generality, take the edge to be $\gamma = \{(x,y) : y = -1, -1 \le x \le 1\}$ of the reference triangle. We construct a basis for the space of polynomials which vanish on $\partial T \setminus \gamma_i$ and express $\widetilde{u}$ in the form

$$\widetilde{u}(x,y) = (1-\xi^2)\left(\frac{1-\eta}{2}\right)^2 \sum_{i+j\le p-2} \widetilde{u}_{ij} P_i^{(2,2)}(\xi)\left(\frac{1-\eta}{2}\right)^i P_j^{(2i+5,0)}(\eta)$$

for suitable coefficients $\{\widetilde{u}_{ij} \in \mathbb{R} : i+j \le p-2\}$, where $\xi = \frac{2(1+x)}{1-y} - 1$ and $\eta = y$. The L^2 norm to minimize can be expressed in terms of $\{\widetilde{u}_{ij}\}$,

$$\|\widetilde{u}\|^2 = \int_{-1}^{1}\int_{-1}^{1} \widetilde{u}^2(x,y)\left(\frac{1-\eta}{2}\right) d\eta d\xi = \sum_{i+j\le p-2} \widetilde{u}_{ij}^2 \mu_i \nu_{ij},$$

where $\nu_{ij} = \int_{-1}^{1}(\frac{1-\eta}{2})^{2i+5} P_j^{(2i+5,0)}(\eta)^2\,d\eta = \frac{1}{i+j+3}$ and μ_i as defined in the lemma statement. The requirement for $\widetilde{u} = u$ on γ means that

$$\widetilde{u}(x,-1) = (1-x^2)\sum_{i+j\le p-2}(-1)^j \widetilde{u}_{ij} P_i^{(2,2)}(x) \implies w_i = \sum_{j=0}^{p-2-i}(-1)^j \widetilde{u}_{ij}.$$

The Cauchy–Schwarz inequality gives

$$w_i^2 \le \left(\sum_{j=0}^{p-2-i} \nu_{ij}^{-1}\right)\left(\sum_{j=0}^{p-2-i} \widetilde{u}_{ij}^2 \nu_{ij}\right) = \frac{1}{2}(p-i-1)(p+i+4)\sum_{j=0}^{p-2-i} \widetilde{u}_{ij}^2 \nu_{ij} \tag{6.4}$$

with equality if there exists a constant λ, such that for all $j \in [0, p-2-i]$ and fixed i, such that $(-1)^j \widetilde{u}_{ij}\nu_{ij}^{1/2} = \lambda \nu_{ij}^{-1/2}$, or equally well, $\widetilde{u}_{ij} = (-1)^j \lambda (i+j+3)$. The choice $\lambda = \frac{w_i}{\sum_{j=0}^{p-2-i} i+j+3}$ gives $w_i = \sum_{j=0}^{p-2-i}(-1)^j \widetilde{u}_{ij}$. Hence, the case of strict equality in (6.4) is achieved.

Direct computation reveals that

$$\|\widetilde{u}\|^2 = \sum_{i=0}^{p-2} \mu_i \sum_{j=0}^{p-2-i} \widetilde{u}_{ij}^2 \nu_{ij} = \sum_{i=0}^{p-2} \frac{\mu_i w_i^2}{\frac{1}{2}(p-i-1)(p+i+4)}$$

and the result follows. ∎

The following discrete weighted Hardy's inequality will prove useful.

Lemma 6.5. *Let $\{v_i\}_{i=0}^p \in \mathbb{R}$ satisfy $\sum_{i=0,\text{even}}^p v_i = 0$ and $\sum_{i=1,\text{odd}}^p v_i = 0$. Then there exists a constant C independent of p such that*

(6.5)
$$\sum_{i=2}^{p} \frac{\widetilde{S}_i^2}{(i-1)^2(2i+1)(i+p+2)(p-i+1)} \le C \sum_{i=0}^{p} \frac{v_i^2}{(2i+1)(i+p+2)(p-i+1)},$$

where

$$\widetilde{S}_i = \begin{cases} |v_0| + |v_2| + \cdots + |v_{i-2}| & \text{if } i \text{ even}, \\ |v_1| + |v_3| + \cdots + |v_{i-2}| & \text{else}. \end{cases} \tag{6.6}$$

Proof. We prove the inequality in the case where all the coefficients with odd indices vanish. Hardy's inequality for weighted sums states that for nonnegative a_k, b_n, c_n,

$$\sum_{n=1}^{\infty} \left(\sum_{k=1}^{n} a_k\right)^2 b_n \le C \sum_{n=1}^{\infty} a_n^2 c_n \tag{6.7}$$

with $C \le 2\sqrt{2}A$ where $A := \sup_{n\in\mathbb{N}} (\sum_{k=n}^{\infty} b_k \sum_{k=1}^{n} c_k^{-1})^{1/2} < \infty$ [18, p. 57]. Choosing $a_k = |v_{2(k-1)}|$ for $k = 1, \dots, \lfloor p/2 \rfloor$ and b_n, c_n for $n = 1, \dots, \lfloor p/2 \rfloor$ to be

$$c_n = \frac{1}{(4n-3)(2n+p)(p-2n+3)},$$
$$b_n = \frac{1}{(2n-1)^2(4n+1)(2n+p+2)(p-2n+1)}$$

with remaining indices chosen to be $a_i, b_i = 0$ and $c_i = 1$ in (6.7) gives the required estimate. A similar argument can be used to obtain the estimate when the coefficients with even indices vanish. The desired estimate then follows by combining the two cases. ∎

The next result gives a bound on the norm of the minimal extension of a polynomial supported on a single edge of a triangle.

LEMMA 6.6. *Let $u \in \mathbb{P}_p(T)$, such that $u(v_i) = 0$ for v_i the vertices of T. Let γ be any edge of T, and let $U \in \mathbb{P}_p(\partial T)$ such that $U|_\gamma = u|_\gamma$ and $U = 0$ on the remaining two edges. Let $\widetilde{U}$ denote the minimal L^2 extension of U; then there exists a constant C independent of p such that*

$$\left\|\widetilde{U}\right\| \le C\|u\|\,.$$

Proof. Without loss of generality, we assume $\gamma = \{(x,y) : y = -1, -1 \le x \le 1\}$ and let Ψ_{ij} be given by (6.1). Since $\{\Psi_{ij}\}_{0\le i,j,i+j\le p}$ forms a basis, we may write $u = \sum_{i+j\le p} u_{ij}\Psi_{ij}$ and denote

$$f = u|_\gamma = \sum_{i+j\le p} (-1)^j u_{ij}\sqrt{\frac{(2i+1)(i+j+1)}{2}}P_i^{(0,0)}(x).$$

Our technique is to express f as a sum of $(1-x^2)P_i^{(2,2)}$, $i = 0,\dots,p-2$, and to then use Lemma 6.4 to calculate $\|\widetilde{U}\|$. Define

$$v_i = \sum_{j=0}^{p-i}(-1)^j u_{ij}\sqrt{\frac{(2i+1)(i+j+1)}{2}}; \tag{6.8}$$

then in order to use Lemma 6.4, we seek coefficients w_i such that

$$f = \sum_{i=0}^{p} v_i P_i^{(0,0)}(x) = (1-x^2)\sum_{i=0}^{p-2} w_i P_i^{(2,2)}(x).$$

Observe that since u vanishes at the vertices of T, we have $u(\pm 1,-1) = 0$, which in turn implies $\sum_{i=0}^{p} v_i = 0$ and $\sum_{i=0}^{p}(-1)^i v_i = 0$, or equally well

$$\sum_{i=0,\text{even}}^{p} v_i = 0, \qquad \sum_{i=1,\text{odd}}^{p} v_i = 0. \tag{6.9}$$

Consequently, we can rewrite f as

$$f = \sum_{i=2,\text{even}}^{p} (P_i^{(0,0)} - P_{i-2}^{(0,0)})S_i + \sum_{i=3,\text{odd}}^{p} (P_i^{(0,0)} - P_{i-2}^{(0,0)})S_i,$$

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.

Using the identity

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

which follows from identities (22.7.15) to (22.7.19) from [1] where P_{n-4} is understood to be 0 for $n < 4$, we have

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

and we deduce that $w_i = \frac{S_{i+4}}{4} - \frac{(i+3)(i+4)}{4(i+1)(i+2)}S_{i+2}$. Writing $S_{i+4} = S_{i+2} - v_{i+2}$, we have

$$w_i = -\frac{v_{i+2}}{4} - \frac{5+2i}{2(i+1)(i+2)}S_{i+2}. \tag{6.10}$$

The Cauchy–Schwarz inequality applied to (6.8) gives

$$v_i^2 \le \sum_{j=0}^{p-i}u_{ij}^2\sum_{j=0}^{p-i}\frac{(2i+1)(i+j+1)}{2} = \frac{(2i+1)(i+p+2)(p-i+1)}{4}\sum_{j=0}^{p-i}u_{ij}^2,$$

which in turn gives

$$\sum_{i=0}^{p}\frac{4v_i^2}{(2i+1)(i+p+2)(p-i+1)} \le \sum_{i=0}^{p}\sum_{j=0}^{p-i}u_{ij}^2 = \|u\|^2. \tag{6.11}$$

Using Lemma 6.4 and the inequality $w_i^2 \le \frac{v_{i+2}^2}{8} + \frac{1}{2}k_i^2S_{i+2}^2$, where $k_i = \frac{5+2i}{2(i+1)(i+2)}$ deduced from (6.10), we have

$$\begin{aligned}\left\|\widetilde{U}\right\|^2 &= \sum_{i=0}^{p-2}\frac{2\mu_iw_i^2}{(p+i+4)(p-i-1)}\\ &\le C\left(\sum_{i=0}^{p-2}\frac{v_{i+2}^2}{(p+i+4)(p-i-1)(2i+5)}+\sum_{i=0}^{p-2}\frac{k_i^2S_{i+2}^2}{(p+i+4)(p-i-1)(2i+5)}\right).\end{aligned}$$

Turning to the first term, thanks to (6.11), we have

$$\sum_{i=0}^{p-2}\frac{v_{i+2}^2}{(p+i+4)(p-i-1)(2i+5)} \le \frac{1}{4}\sum_{i=0}^{p}\frac{4v_i^2}{(2i+1)(i+p+2)(p-i+1)} \le C\|u\|^2.$$

For the second term, we first denote

$$\widetilde{S}_i = \begin{cases}|v_0|+\cdots+|v_{i-2}| & \text{if } i \text{ even},\\ |v_1|+\cdots+|v_{i-2}| & \text{else}\end{cases}$$

so that $S_i^2 \le \widetilde{S}_i^2$. We first note that $k_i \le \frac{2}{i+1}$ and change the index of the summation; then using Lemma 6.5 and (6.11), we obtain

$$\begin{aligned}&\sum_{i=2}^{p}\frac{S_i^2}{(i-1)^2(2i+1)(p+i+2)(p-i+1)}\\ &\le \sum_{i=2}^{p}\frac{\widetilde{S}_i^2}{(i-1)^2(2i+1)(p+i+2)(p-i+1)}\\ &\le C\sum_{i=0}^{p}\frac{v_i^2}{(2i+1)(i+p+2)(p-i+1)} \le C\|u\|^2\end{aligned}$$

and the result follows as claimed. ∎

Finally, we are in a position to give the proof of Theorem 5.5.

Proof. The first step is to construct a suitable decomposition for $u \in X$. Let

$$u_V = \sum_{i=1}^{3} u(v_i)\widetilde{\varphi}_i \in X_V$$

be the interpolant to u at the vertices using the minimal L^2 vertex functions.

Consequently $(u - u_V)|_{\partial T} \in \mathbb{P}_p(\partial T)$ vanishes at the element vertices and can therefore be written in the form

$$u - u_V|_{\partial T} = U_1 + U_2 + U_3,$$

where $U_i \in \mathbb{P}_p(\partial T)$ is supported on edge γ_i. We then let

$$u_{E_i} \in X_{E_i}$$

be the minimal L^2 extension of U_i into the triangle. It follows that

$$u - u_V - \sum_{i=1}^{3} u_{E_i} = u_I \in X_I.$$

Thus $u = u_V + \sum_{i=1}^{3} u_{E_i} + u_I$ is a decomposition of u. It remains to show the decomposition is uniformly bounded.

First, by Lemma 6.1,

$$a_V(u_V, u_V) = \frac{1}{p^4}\sum_{i=1}^{3} u(v_i)^2 \le \frac{3}{p^4} \max_{i\in\{1,2,3\}} u^2(v_i) \le 3C\|u\|^2 . \tag{6.12}$$

For the edge contributions, we use Lemma 6.6 to bound

$$a_{E_i}(u_{E_i}, u_{E_i}) = \|u_{E_i}\|^2 \le C\|u - u_V\|^2 \le 2C\left(\|u\|^2 + \|u_V\|^2\right)$$

and then use the estimate $\|u_V\|^2 \le Ca_V(u_V, u_V)$ from Lemma 5.3 and (6.12) to deduce $\|u_V\|^2 \le \|u\|^2$ and hence $a_{E_i}(u_{E_i}, u_{E_i}) \le C\|u\|^2$.

Finally, as $u_V + \sum_{i=1}^{3} u_{E_i} \in \widetilde{X}_B$, Lemma 5.1 gives us $(u_I, u_V + \sum_{i=1}^{3} u_{E_i}) = 0$, hence

$$a_I(u_I, u_I) = \|u_I\|^2 \le \|u_I\|^2 + \left\|u_V + \sum_{i=1}^{3} u_{E_i}\right\|^2 = \|u\|^2 ,$$

and our result follows. ∎

7. Conclusions. The current work has developed an ASM which results in a uniform condition number in both mesh size h and polynomial order p. The key idea is the construction of a new basis which is used to define the subspace decomposition for the ASM. It is not our intention to suggest that this basis be adopted wholesale, e.g., for Poisson-type problems for which the mass matrix is absent and only the stiffness matrix appears. The key point is that although the spaces used in the description of the ASM are constructed using the specific basis described in section 2, the resulting abstract form of the ASM means that the preconditioner can be applied to whatever basis the reader may care to use through applying a change of basis. For instance, [5] shows how the algorithm can be applied to the Bernstein basis at a cost of $\mathcal{O}(p^3)$ operations. The extension of the ideas to higher dimensions will be the subject of a forthcoming work.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, Dover Books Math. 55, Dover, New York, 1964.

[2] M. Ainsworth, *A preconditioner based on domain decomposition for h-p finite-element approximation on quasi-uniform meshes*, SIAM J. Numer. Anal., 33 (1996), pp. 1358–1376.

[3] 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.

[4] M. Ainsworth and B. Guo, *An additive Schwarz preconditioner for p-version boundary element approximation of the hypersingular operator in three dimensions*, Numer. Math., 85 (2000), pp. 343–366, https://doi.org/10.1007/s002110000134.

[5] M. Ainsworth, S. Jiang, and M. A. Sanchéz, *An $\mathcal{O}(p^3)$ hp-version FEM in two dimensions: Preconditioning and post-processing*, Comput. Methods Appl. Mech. Engrg., to appear.

[6] I. Babuška, A. Craig, J. Mandel, and J. Pitkäranta, *Efficient preconditioning for the p-version finite element method in two dimensions*, SIAM J. Numer. Anal., 28 (1991), pp. 624–661.

[7] I. Babuška and M. Suri, *The p and h-p versions of the finite element method, basic principles and properties*, SIAM Rev., 36 (1994), pp. 578–632.

[8] J. H. Bramble, J. E. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp., 55 (1990), pp. 1–22, https://doi.org/10.2307/2008789.

[9] A. G. Bratsos, *The solution of the two-dimensional sine-Gordon equation using the method of lines*, J. Comput. Appl. Math., 206 (2007), pp. 251–277, https://doi.org/10.1016/j.cam.2006.07.002.

[10] W. Cao and B. Guo, *Preconditioning for the p-version boundary element method in three dimensions with triangular elements*, J. Korean Math. Soc., 41 (2004), pp. 345–368, https://doi.org/10.4134/JKMS.2004.41.2.345.

[11] M. A. Casarin, *Quasi-optimal Schwarz methods for the conforming spectral element discretization*, SIAM J. Numer. Anal., 34 (1997), pp. 2482–2502.

[12] T. F. Chan and T. P. Mathew, *Domain decomposition algorithms*, Acta Numer., 3 (1994), pp. 61–143.

[13] L. Demkowicz, *Computing with hp-Adaptive Finite Elements. Volume* 1: *One and Two Dimensional Elliptic and Maxwell Problems*, Chapman & Hall/CRC Appl. Math. Nonlinear Sci. Ser., Chapman & Hall/CRC, Boca Raton, FL, 2007, https://doi.org/10.1201/9781420011685.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Stud. Math. Sci. 3, Johns Hopkins University Press, Baltimore, 2012.

[15] E. Hairer, *Solving Ordinary Differential Equations* I: *Nonstiff Problems*, Springer, New York, 2011.

[16] E. Hille, G. Szegö, and J. Tamarkin, *On some generalizations of a theorem of A. Markoff*, Duke Math. J., 3 (1937), pp. 729–739.

[17] G. Karniadakis and S. Sherwin, *Spectral/hp Element Methods for Computational Fluid Dynamics*, Oxford University Press, New York, 2013.

[18] A. Kufner, L. Maligranda, and L. Persson, *The Hardy Inequality: About Its History and Some Related Results*, Vydavatelskỳ Servis, 2007, https://books.google.com/books?id=0t6KGQAACAAJ.

[19] 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.

[20] 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, https://doi.org/10.1007/s002110050096.

[21] L. F. Pavarino, *Additive Schwarz methods for the p-version finite element method*, Numer. Math., 66 (1993), pp. 493–515.

[22] C. Schwab, *p- and hp-Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics*, Numer. Math. Sci. Comput., Oxford University Press, New York, 1998.

[23] S. J. Sherwin and G. E. Karniadakis, *A new triangular and tetrahedral basis for high-order (hp) finite element methods*, Internat. J. Numer. Methods Engrg., 38 (1995), pp. 3775–3802.

[24] D. Silvester and A. Wathen, *Fast iterative solution of stabilised Stokes systems.* II. *Using general block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367, https://doi.org/10.1137/0731070.

[25] B. Smith, P. Bjorstad, and W. Gropp, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, Cambridge, UK, 2004.

[26] J. M. Smith, *Efficient Domain Decomposition Preconditioning for the p-Version Finite Element Method: The Mass Matrix.* http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.44.1649.

[27] B. Szabó and I. Babuška, *Finite Element Analysis*, John Wiley & Sons, New York, 1991.

[28] A. Toselli and O. B. Widlund, *Domain Decomposition Methods: Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer, New York, 2005.

[29] 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.