A FAMILY OF RECTANGULAR MIXED ELEMENTS WITH A CONTINUOUS FLUX FOR SECOND ORDER ELLIPTIC PROBLEMS*

TODD ARBOGAST† AND MARY F. WHEELER‡

Abstract. We present a family of mixed finite element spaces for second order elliptic equations in two and three space dimensions. Our spaces approximate the vector flux by a continuous function. Our spaces generalize certain spaces used for approximation of Stokes problems. The finite element method incorporates projections of the Dirichlet data and certain low order terms. The method is locally conservative on the average. Suboptimal convergence is proven and demonstrated numerically. The key result is to construct a flux π-projection operator that is bounded in the Sobolev space H^1, preserves a projection of the divergence, and approximates optimally. Moreover, the corresponding Raviart–Thomas flux preserving π-projection operator is an L^2-projection when restricted to this family of spaces.

1. Introduction. Mixed finite element methods have been used effectively to solve many problems, including second order elliptic problems [10, 15, 31, 34]. Both the scalar variable and its vector flux are approximated directly. While it is necessary to approximate the flux in $H(\mathrm{div})$, the space of L^2 vectors whose divergence is also in L^2, it is not necessary that the flux be fully continuous. In the usual mixed spaces (see, e.g., [12, 13, 14, 17, 29, 31]), only the normal component of the approximate flux is continuous. The tangential components are discontinuous across element boundaries.

In some applications, it is desirable that the flux be continuous. The applications we have in mind come from simulating fluid flow in a porous medium [7, 30, 32]. Miscible displacement in a petroleum reservoir or groundwater transport problems require the solution to a system of equations in model form:

$$(1.1) \qquad ap + \nabla \cdot \mathbf{u} = b, \quad \Omega,$$

$$(1.2) \qquad d\mathbf{u} = -\nabla p + \mathbf{c}, \quad \Omega,$$

$$(1.3) \qquad \phi \frac{\partial c}{\partial t} - \nabla \cdot \big(D(\mathbf{u})\nabla c - \mathbf{u}c\big) = \hat{c}f^+ - cf^-, \quad \Omega, \ t > 0,$$

where p is the pressure, $\mathbf{u}$ is the Darcy velocity (i.e., the flux), c is the concentration of a dissolved chemical that is transported by the flow, and a, b, $\mathbf{c}$, d, ϕ, D, and $\hat{c}$ are

† Institute for Computational Engineering and Sciences, University of Texas, 1 University Station C0200, Austin, TX 78712, and Mathematics Department, University of Texas, 1 University Station C1200, Austin, TX 78712–0257 (arbogast@ices.utexas.edu). This author was supported in part by U.S. National Science Foundation grant DMS-0074310.

‡Institute for Computational Engineering and Sciences, University of Texas, 1 University Station C0200, Austin, TX 78712, and Department of Aerospace Engineering & Engineering Mechanics, University of Texas, 1 University Station C0600, Austin, TX 78712, and Department of Petroleum & Geosystems Engineering, University of Texas, 1 University Station C0300, Austin, TX 78712 (mfw@ices.utexas.edu).

various parameters. To these equations, we must add boundary and initial conditions. For the subsystem (1.1)–(1.2), let $\partial\Omega$ be decomposed into Γ_N and Γ_D, and set

$$(1.4) \qquad\qquad\qquad \mathbf{u} \cdot \nu = g_N, \quad \Gamma_N,$$

$$(1.5) \qquad\qquad\qquad p = p_D, \quad \Gamma_D,$$

where ν is the outer unit normal vector, g_N is the given boundary flux, and p_D is the given boundary pressure.

Since $D(\mathbf{u}) \approx 0$, equation (1.3) is nearly hyperbolic. Characteristic methods have been successful in treating this equation (see, e.g., [4, 20, 22, 23]); however, they rely heavily on the velocity $\mathbf{u}$. To obtain good characteristic trace-backs, $\mathbf{u}_h \approx \mathbf{u}$ should satisfy

(i) $\nabla \cdot \mathbf{u}_h = \mathcal{P}_W f$, where $\mathcal{P}_W$ is an appropriate projection;

(ii) $\mathbf{u}_h$ is continuous.

Property (i) gives a proper divergence to the flow so that mass is conserved, while (ii) is required for consistency in tracing regions through the flow field.

A second application involves the coupling of Stokes flow with Darcy flow in a region with open channel flow adjacent to a porous medium [8]. Usual approaches require that the Stokes equations be approximated by a continuous velocity, since it must remain in H^1. However, this is not properly matched on the open/porous interface to a discontinuous Darcy velocity. A continuous Darcy velocity would therefore be desirable. These spaces have been exploited in [1, 2, 3].

Current mixed methods achieve (i) at the expense of (ii). Our goal is to relax (i), so that it holds only "on the average", but maintain (ii). Brezzi, Fortin, and Marini [16] presented a stabilization technique that allows the use of continuous finite element spaces. Their technique involves a modification of the usual mixed equations. Herein, we provide a family of mixed methods that is stable for the original set of mixed equations. These methods are defined on rectangular grids, and they generalize the Stokes elements of Fortin [24] (cf. Bernardi and Raugel [9]). We present the full development for two dimensions, and discuss the three dimensional case in the last section.

2. Some general notation. Throughout the paper, for domain ω, we denote by $L^p(\omega)$ the usual Lebesgue space of index p, $1 \le p \le \infty$, and by $W^{k,p}(\omega)$ the usual Sobolev space of k weak derivatives in $L^p(\omega)$. We denote by $(\cdot, \cdot)_\omega$ the $L^2(\omega)$ inner product (i.e., Lebesgue integration over ω). Moreover, $\| \cdot \|_{k,\omega}$ is the norm of $H^k(\omega) \equiv W^{k,2}(\omega)$, and $| \cdot |_{r,\omega}$ denotes the $H^r(\omega)$ seminorm. In the notation we may suppress ω when it is Ω. On a domain boundary e, we use the notation $\langle \cdot, \cdot \rangle_e$ for the $L^2(e)$ inner product. For d-dimensional set S, $|S|$ is its d-dimensional Lebesgue measure.

Let $P_k(\omega)$ denote the space of polynomials of degree at most k over the set ω. Moreover, in $\mathbb{R}^2$, let $Q_{i,j}(\omega)$ be the set of polynomials of degree at most i in x and j in y over ω, and similarly define $Q_{i,j,k}(\omega)$ in $\mathbb{R}^3$. We will make use of scaling arguments, so let us define $\hat{R} = [-1, 1]^2$ as our reference rectangle. Moreover, let $\hat{\lambda}_j$ denote the Legendre polynomial of degree j on $[-1, 1]$. Recall that they are $L^2([-1, 1])$-orthogonal polynomials, and that by convention, they are normalized so that $\hat{\lambda}_j(1) = 1$; then also $\hat{\lambda}_j(-1) = (-1)^j$.

Finally, if X is a closed subspace of $L^2(\omega)$, we denote by $\mathcal{P}_X : L^2(\omega) \to X$ the $L^2(\omega)$-projection operator onto X defined for $\psi \in L^2(\omega)$ as the unique $\mathcal{P}_X \psi \in X$ such that

$$(\psi - \mathcal{P}_X \psi, \varphi)_\omega = 0 \quad \forall \varphi \in X.$$

3. An illustrative example. To illustrate our finite element spaces, consider (1.1)–(1.2) in mixed form for $\Omega \subset \mathbb{R}^2$ a bounded domain, with the natural boundary conditions $\Gamma_N = \partial\Omega$ and $g_N = 0$, and a and $\mathbf{c}$ set to 0:

$$(3.1) \qquad\qquad (\nabla \cdot \mathbf{u}, w) = (b, w) \qquad \forall w \in L^2(\Omega)/\mathbb{R},$$

$$(3.2) \qquad\qquad (d\mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) = 0 \quad \forall \mathbf{v} \in H_0(\text{div}; \Omega),$$

where $H_0(\text{div}; \Omega)$ is the subset of $H(\text{div}; \Omega)$ with vanishing normal trace on $\partial\Omega$. Let $\mathbf{V}_h \times W_h$ denote some mixed finite element space, and solve (3.1)–(3.2) for $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$ with the restrictions that $w \in W_h$ and $\mathbf{v} \in \mathbf{V}_h$.

If $\mathbf{V}_h \times W_h$ is the lowest order Raviart–Thomas space [29, 31, 34], the solution $\mathbf{u}_h$ is discontinuous. On a rectangular element R, the x-coordinate of $\mathbf{u}_h$, $u_{h,1}$, is in

$$V_{h,1}^{RT_0}(R) = Q_{1,0}(R).$$

This space has one degree of freedom for each edge normal to the x-direction. If we add four degrees of freedom by defining

$$V_{h,1}(R) = Q_{1,2}(R),$$

the extra corner degrees of freedom allow us to enforce continuity, while the two edge degrees of freedom allow us to maintain the *average* flux across each normal edge, i.e., the proper average divergence of the flow. A similar modification in y, $V_{h,2}(R) = Q_{2,1}(R)$, gives a new element with the required properties on all edges of R. This is a Stokes element due to Fortin [24].

4. The spaces in two dimensions on rectangles. Assume $\Omega \subset \mathbb{R}^2$. Let $\mathcal{T}_h$ denote a quasi-regular, conforming finite element partition of Ω into rectangles of diameter bounded by h. By quasi-regular, we mean that the aspect ratio of the rectangles is bounded by a fixed constant. We now define our family of mixed spaces; it contains the element of the previous section as its lowest order member.

DEFINITION 4.1. *For each integer $k \geq 1$, the linear space $\mathbf{V}_h^k \times W_h^{k-1}$ is defined so that, for any rectangle $R \in \mathcal{T}_h$,*

$$\mathbf{V}_h^k(R) = Q_{k,k+1}(R) \times Q_{k+1,k}(R),$$
$$W_h^{k-1}(R) = Q_{k-1,k-1}(R),$$

and

$$\mathbf{V}_h^k = \{\mathbf{v} \in (C^0(\Omega))^2 : \mathbf{v}|_R \in \mathbf{V}_h^k(R) \quad \forall R \in \mathcal{T}_h\},$$
$$W_h^{k-1} = \{w \in L^2(\Omega) : w|_R \in W_h^{k-1}(R) \quad \forall R \in \mathcal{T}_h\}.$$

A local basis for $W_h^{k-1}(R)$ is trivial to construct; moreover, since W_h^{k-1} is discontinuous across element boundaries, a global basis can be constructed immediately.

A local nodal basis can be defined for $\mathbf{V}_h^k(R)$ by the degrees of freedom given in the next lemma. Moreover, because these degrees of freedom uniquely determine the function on each edge, they can be pieced together across edges and vertices to form a global basis for the continuous function space $\mathbf{V}_h^k$.

LEMMA 4.2. *For R a rectangle, $\mathbf{u}_h \in \mathbf{V}_h^k(R)$ is uniquely defined by the following degrees of freedom:*

(i) *for each corner point $P \in \partial R$ and Cartesian direction $j = 1, 2$,*

$$DOF_{P,j}^{(i)}(\mathbf{u}_h) = u_{h,j}(P) = \mathbf{u}_h(P) \cdot \mathbf{e}_j;$$

(ii) *on each edge $e \subset \partial R$,*

$$DOF_{e,\lambda}^{(ii)}(\mathbf{u}_h) = \langle \mathbf{u}_h \cdot \tau_e, \lambda \rangle_e \quad \forall \lambda \in P_{k-2}(e),$$

where τ_e is a unit tangential direction;
(iii) *on each edge $e \subset \partial R$,*

$$DOF_{e,\lambda}^{(iii)}(\mathbf{u}_h) = \langle \mathbf{u}_h \cdot \nu_e, \lambda \rangle_e \quad \forall \lambda \in P_{k-1}(e),$$

where ν_e is the outer unit normal direction;
(iv) *over R,*

$$DOF_{\mathbf{v}}^{(iv)}(\mathbf{u}_h) = (\mathbf{u}_h, \mathbf{v})_R \quad \forall \mathbf{v} \in Q_{k-2,k-1}(R) \times Q_{k-1,k-2}(R).$$

Moreover, on each edge $e \subset \partial R$, $\mathbf{u}_h|_e$ is uniquely defined by the degrees of freedom (i)–(iii) restricted to e.

Proof. We restrict our analysis to the case where $R = \hat{R} = [-1, 1]^2$; an affine map can be used to show the result for a general rectangle R. As usual, since our function spaces are finite-dimensional vector spaces, the degrees of freedom uniquely determine the function if and only if both the dimension of the function space and the number of independent degrees of freedom agree; whenever the degrees of freedom vanish, the function also vanishes.

We begin by considering $\mathbf{u}_h$ on an edge e. Now $\mathbf{u}_h \cdot \tau_e \in P_k(e)$ and $\mathbf{u}_h \cdot \nu_e \in P_{k+1}(e)$, so the total dimension of this space is $2k + 3$. The number of degrees of freedom that act on e are 4 for (i), $k - 1$ for (ii), k for (iii), and 0 for (iv), leading to the same number $2k + 3$ degrees of freedom on e. Suppose that the degrees of freedom (i)–(iii) restricted to e vanish. We conclude from (i) that

$$\mathbf{u}_h \cdot \tau_e(\xi) = (1 - \xi^2) \, q(\xi),$$

where $q \in P_{k-2}$ (and $\mathbf{u}_h \cdot \tau_e \equiv 0$ if $k = 1$). Then (ii) implies that $q = 0$, and we conclude that $\mathbf{u}_h \cdot \tau_e \equiv 0$ on e. Similarly, we conclude from (i) and (iii) that $\mathbf{u}_h \cdot \nu_e \equiv 0$ on e. We have thereby demonstrated the last statement of the lemma.

Note that the total number of degrees of freedom is 8 for (i), $4(k - 1)$ for (ii), $4k$ for (iii), and $2(k - 1)k$ for (iv), so that the total is $2k^2 + 6k + 4$. This is the same as

$$\dim \mathbf{V}_h^k(R) = 2(k + 1)(k + 2) = 2k^2 + 6k + 4.$$

So suppose that all degrees of freedom of $\mathbf{u}_h$ vanish. We have already shown that then $\mathbf{u}_h|_{\partial R}$ vanishes, so

$$\mathbf{u}_h(x, y) = (1 - x^2)(1 - y^2) \, \mathbf{v}(x, y)$$

for some $\mathbf{v}(x, y) \in Q_{k-2,k-1}(R) \times Q_{k-1,k-2}(R)$. By degree of freedom (iv), we conclude that $\mathbf{v} \equiv 0$, and so also $\mathbf{u}_h \equiv 0$, completing the proof. $\square$

For completeness and future reference, we show how to construct an explicit nodal basis on the reference rectangle $\hat{R} = [-1, 1]^2$. A nodal basis has the property that each member has one degree of freedom evaluate to 1 and the rest to 0. Recall that $\hat{\lambda}_j$

is the Legendre polynomial of degree j on $\hat{R}$, that $\langle \hat{\lambda}_i, \hat{\lambda}_j \rangle_{[-1,1]} = 0$ if $i \neq j$, and that $\hat{\lambda}_j(1) = 1$ and $\hat{\lambda}_j(-1) = (-1)^j$. For degrees of freedom (ii)–(iv), we need to select a basis for the test spaces. Since $\{\hat{\lambda}_j\}_{j \leq k}$ forms a basis for $P_k(-1,1)$, we can restrict the polynomials in the degrees of freedom (ii) and (iii) to Legendre polynomials, and to tensor products of such in (iv).

For degree of freedom (i), we define for corner point $P = (-1, -1)$ and direction 1 the basis function

$$\hat{\mathbf{v}}_{P,1}^{(i)} = -\frac{1}{4}(\hat{\lambda}_k(\hat{x}) - \hat{\lambda}_{k-1}(\hat{x}))(\hat{\lambda}_{k+1}(\hat{y}) - \hat{\lambda}_k(\hat{y}))\mathbf{e}_1.$$

This function has the property that it is $\mathbf{e}_1$ at P and vanishes at the three other corner points, and the degrees of freedom (ii)–(iv) vanish. We similarly define a basis function for the other three corner points, and for direction 2. For degree of freedom (ii), with edge $e = (-1, 1) \times \{-1\}$ and $\lambda = \hat{\lambda}_j(\hat{x})$, $0 \leq j \leq k-2$, we define

$$\hat{\mathbf{v}}_{e,j}^{(ii)} = \alpha_{e,j}^{(ii)}(\hat{\lambda}_j(\hat{x}) - \hat{\lambda}_{k_j}(\hat{x}))(\hat{\lambda}_{k+1}(\hat{y}) - \hat{\lambda}_k(\hat{y}))\mathbf{e}_1,$$

where k_j is either $k-1$ or k so that k_j and j have the same even/odd parity (i.e., so that $\hat{\lambda}_{k_j}(-1) = \hat{\lambda}_j(-1)$), and $\alpha_{e,j}^{(ii)}$ is chosen to fix the normalization

$$\mathrm{DOF}_{e,\lambda_j}^{(ii)}(\hat{\mathbf{v}}_{e,j}^{(ii)}) = \langle \hat{\mathbf{v}}_{e,j}^{(ii)} \cdot \tau_e, \hat{\lambda}_j \rangle_e = 1.$$

The other basis functions of type (ii) are defined similarly. For degree of freedom (iii), with edge $e = \{-1\} \times (-1, 1)$ and $\lambda = \hat{\lambda}_j(\hat{y})$, $0 \leq j \leq k-1$, we define

$$\hat{\mathbf{v}}_{e,j}^{(iii)} = \alpha_{e,j}^{(iii)}(\hat{\lambda}_k(\hat{x}) - \hat{\lambda}_{k-1}(\hat{x}))(\hat{\lambda}_j(\hat{y}) - \hat{\lambda}_{k_j}(\hat{y}))\mathbf{e}_1,$$

where k_j is either k or $k+1$ so that k_j and j have the same parity and $\alpha_{e,j}^{(iii)}$ is chosen to make

$$\mathrm{DOF}_{e,\lambda_j}^{(iii)}(\hat{\mathbf{v}}_{e,j}^{(iii)}) = \langle \hat{\mathbf{v}}_{e,j}^{(iii)} \cdot \nu_e, \hat{\lambda}_j \rangle_e = 1.$$

The other basis functions of type (iii) are defined similarly. Finally, for degree of freedom (iv), with $\mathbf{v} = \hat{\lambda}_j(\hat{x})\hat{\lambda}_\ell(\hat{y})\mathbf{e}_1$, $0 \leq j \leq k-2$, $0 \leq \ell \leq k-1$, we define

$$\hat{\mathbf{v}}_{j,\ell}^{(iv)} = \alpha_{j,\ell}^{(iv)}(\hat{\lambda}_j(\hat{x}) - \hat{\lambda}_{k_j}(\hat{x}))(\hat{\lambda}_\ell(\hat{y}) - \hat{\lambda}_{k_\ell}(\hat{y}))\mathbf{e}_1,$$

where k_j is either $k-1$ or k so that $\hat{\lambda}_{k_j}(-1) = \hat{\lambda}_j(-1)$, and k_ℓ is either k or $k+1$ so that $\hat{\lambda}_{k_\ell}(-1) = \hat{\lambda}_\ell(-1)$, and $\alpha_{j,\ell}^{(iv)}$ is defined so that

$$\mathrm{DOF}_{\hat{\lambda}_j \hat{\lambda}_\ell}^{(iv)}(\hat{\mathbf{v}}_{j,\ell}^{(iv)}) = (\hat{\mathbf{v}}_{j,\ell}^{(ii)}, \hat{\lambda}_j \hat{\lambda}_\ell)_{\hat{R}} = 1.$$

The other type (iv) nodal basis functions are defined similarly.

The nodal basis on $\mathcal{T}_h$ is constructed from these local basis functions via local affine mappings. On $R \in \mathcal{T}_h$, we would map R to $\hat{R}$ by $x \mapsto \hat{x} \equiv (x - x_0)/h_1$ and $y \mapsto \hat{y} \equiv (y - y_0)/h_2$, where R has side lengths h_1 and h_2 and lower left corner (x_0, y_0). The important part of the construction is that the reference basis functions $\hat{\mathbf{v}}^{(i)}, \ldots,$ $\hat{\mathbf{v}}^{(iv)}$ are independent of h.

We close this section with a remark about the finite element basis. We chose degrees of freedom and a local basis that are useful for the numerical analysis that

follows. This is primarily due to degree of freedom (iii), which says that moments of the normal flux are controlled. However, an equivalent set of degrees of freedom would be to replace (ii) and (iii) by evaluation at an appropriate number of points along the boundary. This would be a better basis for implementation, since it is simpler to construct.

5. A π operator. For simplicity, let $\mathcal{P}_W^{k-1} : L^2(\Omega) \to W_h^{k-1}$ denote the $L^2(\Omega)$-projection $\mathcal{P}_{W_h^{k-1}}$. As is usual for mixed spaces, we will define a π operator [15, 21, 31] for our spaces. Our π operator should map $(H^1(\Omega))^2$ onto $\mathbf{V}_h^k$. When $\mathbf{u} \in (H^1(\Omega))^2$, the normal or tangential trace of $\mathbf{u}$ on an edge e is in $H^{1/2}(e)$, and so degrees of freedom (ii) and (iii), and also (iv), are defined. However, degree of freedom (i) causes some problems.

We can resolve the difficulty by using the Clément [19] (or the Scott–Zhang [33]) interpolant $\mathcal{I}^k : H^1(\Omega) \to Q_h^k$, where Q_h^k is the space of continuous functions with $Q_h^k|_R = Q_{k,k}(R)$ for any $R \in \mathcal{T}_h$. For completeness, we define $\mathcal{I}^k$ as in [19]. The interpolant of $\psi \in L^2(\Omega)$ is defined at the nodal points of Q_h^k by setting the value to a local L^2-projection of ψ. Let N denote the set of nodal points of Q_h^k; these are, for example, the union over $R \in \mathcal{T}_h$ of the $(k+1)^2$ grid points of a $k \times k$ uniform grid over R. For each $P \in N$, let

$$\Delta_P = \bigcup_{R \in \mathcal{T}_h \text{ with } P \in R} R$$

be the union of the rectangles containing P. Then define $\psi_P \in Q_{k,k}(\Delta_P)$ by

$$(\psi - \psi_P, \varphi)_{\Delta_P} = 0 \quad \forall \varphi \in Q_{k,k}(\Delta_P),$$

and set $\mathcal{I}^k \psi(P) = \psi_P(P)$. This uniquely defines $\mathcal{I}^k \psi \in Q_h^k$; moreover, we have the estimate

$$(5.1) \qquad \|\psi - \mathcal{I}^k \psi\|_j \le C \|\psi\|_r h^{r-j}, \quad j \le r \le k+1, \ j = 0, 1.$$

DEFINITION 5.1. *For $\mathbf{u} \in (H^1(\Omega))^2$, let $\pi^k \mathbf{u} \in \mathbf{V}_h^k$ be defined as the interpolant of the degrees of freedom from Lemma 4.2, modified by the Clément operator $\mathcal{I}^k$. That is, for each $R \in \mathcal{T}_h$, we require the following:*
  (1) *For each corner point $P \in \partial R$ and direction $j = 1, 2$,*

$$\pi^k \mathbf{u}(P) \cdot \mathbf{e}_j = \mathcal{I}^k \mathbf{u}(P) \cdot \mathbf{e}_j.$$

  (2) *On each edge $e \subset \partial R$,*

$$\langle \pi^k \mathbf{u} \cdot \tau_e, \lambda \rangle_e = \langle \mathbf{u} \cdot \tau_e, \lambda \rangle_e \quad \forall \lambda \in P_{k-2}(e).$$

  (3) *On each edge $e \subset \partial R$,*

$$\langle \pi^k \mathbf{u} \cdot \nu_e, \lambda \rangle_e = \langle \mathbf{u} \cdot \nu_e, \lambda \rangle_e \quad \forall \lambda \in P_{k-1}(e).$$

  (4) *Over R,*

$$(\pi^k \mathbf{u}, \mathbf{v})_R = (\mathbf{u}, \mathbf{v})_R \quad \forall \mathbf{v} \in Q_{k-2,k-1}(R) \times Q_{k-1,k-2}(R).$$

The operator π^k is linear, and it is well defined by Lemma 4.2. Before deriving properties of this operator, we consider its explicit construction using the global nodal basis given in section 4. Let us denote this basis as $\{\mathbf{v}_{j_1}^{(\mathrm{i})}, \mathbf{v}_{j_2}^{(\mathrm{ii})}, \mathbf{v}_{j_3}^{(\mathrm{iii})}, \mathbf{v}_{j_4}^{(\mathrm{iv})}\}_{j_1, j_2, j_3, j_4}$, where the superscript designates the degree of freedom type, and for consistency of notation below, the index ranges do not overlap. Basis function $\mathbf{v}_{j_2}^{(\mathrm{ii})}$, of degree of freedom type (ii), is defined with respect to some grid edge e_{j_2} and Legendre polynomial λ^{j_2} on e_{j_2}, and similarly for $\mathbf{v}_{j_3}^{(\mathrm{iii})}$. For $\mathbf{v}_{j_4}^{(\mathrm{iv})}$, it is defined with respect to rectangle R_{j_4} and a tensor product of Legendre polynomials $\mathbf{v}^{j_4}$.

We claim that we can represent, for $\mathbf{u} \in (H^1(\Omega))^2$,

$$\pi^k \mathbf{u} = \mathcal{I}^k \mathbf{u} + \sum_{j_2} \frac{1}{|e_{j_2}|} \langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \tau_{e_{j_2}}, \lambda^{j_2} \rangle_{e_{j_2}} \, \mathbf{v}_{j_2}^{(\mathrm{ii})}$$

$$+ \sum_{j_3} \frac{1}{|e_{j_3}|} \langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \nu_{e_{j_3}}, \lambda^{j_3} \rangle_{e_{j_3}} \, \mathbf{v}_{j_3}^{(\mathrm{iii})}$$

(5.2)
$$+ \sum_{j_4} \frac{1}{|R_{j_4}|} (\mathbf{u} - \mathcal{I}^k \mathbf{u}, \mathbf{v}^{j_4}) \, \mathbf{v}_{j_4}^{(\mathrm{iv})},$$

which is indeed in $\mathbf{V}_h^k$. We note that, after a local change of variables,

$$\langle \mathbf{v} \cdot \xi, \lambda \rangle_e = |e| \langle \hat{\mathbf{v}} \cdot \hat{\xi}, \hat{\lambda} \rangle_{\hat{e}} \quad \text{and} \quad (\mathbf{v}, \psi)_R = |R| (\hat{\mathbf{v}}, \hat{\psi})_{\hat{R}},$$

and thus by our normalization of the reference basis functions, the degrees of freedom (ii)–(iv) of $\pi^k \mathbf{u}$ match $\mathbf{u}$, and $\pi^k \mathbf{u}$ has the correct values at the grid points.

LEMMA 5.2. *Assume that $u \in (H^1(\Omega))^2$.*

(a) *The linear operator π^k is bounded on $(H^1(\Omega))^2$ independently of h.*

(b) *There exists some constant C independent of h such that for $R \in \mathcal{T}_h$ and $\mathbf{u} \in (H^r(\Delta_R))^2$,*

$$|\pi^k \mathbf{u} - \mathbf{u}|_{j,R} \le C |\mathbf{u}|_{r, \Delta_R} h_R^{r-j}, \quad 1 \le r \le k+1, \; j = 0, 1,$$

where $h_R = \operatorname{diam}(R)$ and Δ_R is the union of R and its nearest neighboring elements in $\mathcal{T}_h$.

(c) *For $\mathbf{u} \in (H^r(\Omega))^2$,*

$$|\pi^k \mathbf{u} - \mathbf{u}|_j \le C |\mathbf{u}|_r h^{r-j}, \quad 1 \le r \le k+1, \; j = 0, 1.$$

(d) $\mathcal{P}_W^{k-1} \nabla \cdot \mathbf{u} = \mathcal{P}_W^{k-1} \nabla \cdot \pi^k \mathbf{u}$.

Proof. To show (a), we use the representation (5.2) derived above, which when squared implies

$$|\pi^k \mathbf{u}|^2 \le |\mathcal{I}^k \mathbf{u}|^2 + \sum_{j_2} \frac{1}{|e_{j_2}|^2} |\langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \tau_{e_{j_2}}, \lambda^{j_2} \rangle_{e_{j_2}} \, \mathbf{v}_{j_2}^{(\mathrm{ii})}|^2$$

$$+ \sum_{j_3} \frac{1}{|e_{j_3}|^2} |\langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \nu_{e_{j_3}}, \lambda^{j_3} \rangle_{e_{j_3}} \, \mathbf{v}_{j_3}^{(\mathrm{iii})}|^2$$

$$+ \sum_{j_4} \frac{1}{|R_{j_4}|^2} |(\mathbf{u} - \mathcal{I}^k \mathbf{u}, \mathbf{v}^{j_4}) \, \mathbf{v}_{j_4}^{(\mathrm{iv})}|^2,$$

since at each point $x \in \Omega$ the sums are finite with the number depending on k but not on h. A similar expression holds for the gradient, and so, after integrating,

$$\|\pi^k \mathbf{u}\|_1^2 \leq \|\mathcal{I}^k \mathbf{u}\|_1^2 + \sum_{j_2} \frac{1}{|e_{j_2}|^2} \|\langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \tau_{e_{j_2}}, \lambda^{j_2} \rangle_{e_{j_2}}^2 \, \mathbf{v}_{j_2}^{(ii)}\|_1^2$$

$$+ \sum_{j_3} \frac{1}{|e_{j_3}|^2} \|\langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \nu_{e_{j_3}}, \lambda^{j_3} \rangle_{e_{j_3}}^2 \, \mathbf{v}_{j_3}^{(iii)}\|_1^2$$

(5.3)
$$+ \sum_{j_4} \frac{1}{|R_{j_4}|^2} \|(\mathbf{u} - \mathcal{I}^k \mathbf{u}, \mathbf{v}^{j_4})^2 \, \mathbf{v}_{j_4}^{(iv)}\|_1^2.$$

We begin with the second term on the right-hand side above. By the standard affine change of variables $R \mapsto \hat{R}$ and the quasi regularity of the grid, we deduce that for any ψ

(5.4) $$\|\psi\|_{0,R} = |R|^{1/2} \|\hat{\psi}\|_{0,\hat{R}},$$

(5.5) $$|\psi|_{1,R} \leq C|R|^{1/2} h^{-1} |\hat{\psi}|_{1,\hat{R}} \leq C |\hat{\psi}|_{1,\hat{R}} \leq C |\psi|_{1,R},$$

since in two dimensions $|R|^{1/2} = O(h)$. Now

$$\|\mathbf{v}_{j_2}^{(ii)}\|_1 = \left\{ \sum_R \|\mathbf{v}_{j_2}^{(ii)}\|_{1,R}^2 \right\}^{1/2} \leq C \left\{ \sum_R \|\hat{\mathbf{v}}_{j_2}^{(ii)}\|_{1,\hat{R}}^2 \right\}^{1/2} \leq C,$$

since $\hat{\mathbf{v}}_{j_2}^{(ii)}$ is supported in at most four rectangles and is independent of h. Moreover, if $e_{j_2} \subset R' \in \mathcal{T}_h$, then by the trace theorem (see, e.g., [26]),

$$\frac{1}{|e_{j_2}|} |\langle \mathbf{v} \cdot \tau_{e_{j_2}}, \lambda^{j_2} \rangle_{e_{j_2}}| = |\langle \hat{\mathbf{v}} \cdot \hat{\tau}_{\hat{e}_{j_2}}, \hat{\lambda}^{j_2} \rangle_{\hat{e}_{j_2}}|$$

$$\leq C \|\hat{\mathbf{v}}\|_{1,\hat{R}'} \|\hat{\lambda}^{j_2}\|_{0,\hat{e}_{j_2}} \leq C \{ h^{-1} \|\mathbf{v}\|_{0,R'} + |\mathbf{v}|_{1,R'} \},$$

and so

$$\sum_{j_2} \frac{1}{|e_{j_2}|^2} \|\langle \mathbf{v} \cdot \tau_{e_{j_2}}, \lambda^{j_2} \rangle_{e_{j_2}} \mathbf{v}_{j_2}^{(ii)}\|_1^2 \leq C \sum_{j_2} \{ h^{-1} \|\mathbf{v}\|_{0,R'} + |\mathbf{v}|_{1,R'} \}^2$$

(5.6)
$$\leq C \{ h^{-2} \|\mathbf{v}\|_0^2 + |\mathbf{v}|_1^2 \},$$

since a given rectangle $R' \supset e_{j_2}$ appears at most four times in the sum. Similarly,

(5.7) $$\sum_{j_3} \frac{1}{|e_{j_3}|^2} \|\langle \mathbf{v} \cdot \nu_{e_{j_3}}, \lambda^{j_3} \rangle_{e_{j_3}} \mathbf{v}_{j_3}^{(iii)}\|_1^2 \leq C \{ h^{-2} \|\mathbf{v}\|_0^2 + |\mathbf{v}|_1^2 \}$$

and

(5.8) $$\sum_{j_4} \frac{1}{|R_{j_4}|^2} \|(\mathbf{v}, \mathbf{v}^{j_4}) \, \mathbf{v}_{j_4}^{(iv)}\|_1^2 \leq C h^{-2} \|\mathbf{v}\|_0^2 .$$

Thus from (5.3) we deduce that

(5.9) $$\|\pi^k \mathbf{u}\|_1 \leq C \{ \|\mathcal{I}^k \mathbf{u}\|_1 + h^{-1} \|\mathbf{u} - \mathcal{I}^k \mathbf{u}\|_0 + |\mathbf{u} - \mathcal{I}^k \mathbf{u}|_1 \} \leq C \|\mathbf{u}\|_1.$$

Now (b), with $j = 1$, follows from the Bramble–Hilbert lemma [10], since locally π^k is a kth degree polynomial preserving operator. To prove the $j = 0$ case when $k > 1$, we need only compute

$$\|\mathbf{u} - \pi^k \mathbf{u}\|_{0,R}^2 = \|\mathbf{u} - \bar{\mathbf{u}} - (\pi^k \mathbf{u} - \overline{\pi^k \mathbf{u}})\|_{0,R}^2 \leq C \|\mathbf{u} - \pi^k \mathbf{u}\|_{1,R}^2 h^2,$$

where the over-line denotes the local average and we use the fact that $\bar{\mathbf{u}} = \overline{\pi^k \mathbf{u}}$ by degree of freedom (iv). The case of $k = 1$ follows directly from a careful scaling analysis of (5.2) as above and (5.1). The argument appears in [26]: noting (5.4) and modifying (5.6)–(5.8), we see as in (5.9) that

$$\|\mathbf{u} - \pi^k \mathbf{u}\|_0 \leq \|\mathbf{u} - \mathcal{I}^k \mathbf{u}\|_0 + \sum_{j_2} \frac{1}{|e_{j_2}|} \|\langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \tau_{e_{j_2}}, \lambda^{j_2}\rangle_{e_{j_2}} \mathbf{v}_{j_2}^{(ii)}\|_0$$

$$+ \sum_{j_3} \frac{1}{|e_{j_3}|} \|\langle (\mathbf{u} - \mathcal{I}^k \mathbf{u}) \cdot \nu_{e_{j_3}}, \lambda^{j_3}\rangle_{e_{j_3}} \mathbf{v}_{j_3}^{(iii)}\|_0$$

$$+ \sum_{j_4} \frac{1}{|R_{j_4}|} \|(\mathbf{u} - \mathcal{I}^k \mathbf{u}, \mathbf{v}^{j_4}) \mathbf{v}_{j_4}^{(iv)}\|_0$$

$$\leq C\{\|\mathbf{u} - \mathcal{I}^k \mathbf{u}\|_0 + h|\mathbf{u} - \mathcal{I}^k \mathbf{u}|_1\} \leq C h^r \|\mathbf{u}\|_r,$$

where $r = 1$ or 2.

Result (c) follows from (b). For (d), let $w \in W_h^{k-1}$ and compute

$$(\nabla \cdot \mathbf{u}, w)_R = \langle \mathbf{u} \cdot \nu, w\rangle_{\partial R} - (\mathbf{u}, \nabla w)_R$$
$$= \langle \pi^k \mathbf{u} \cdot \nu, w\rangle_{\partial R} - (\pi^k \mathbf{u}, \nabla w)_R = (\nabla \cdot \pi^k \mathbf{u}, w)_R,$$

since $w|_e \in P_{k-1}(e)$ and $\nabla w|_R \in Q_{k-2,k-1}(R) \times Q_{k-1,k-2}(R)$.   □

Remark 5.1. The usual mixed spaces satisfy

$$(5.10) \qquad\qquad \mathcal{P}_{W_{h,usual}} \nabla \cdot \mathbf{u} = \nabla \cdot \pi_{usual} u,$$

since $\nabla \cdot \mathbf{V}_{h,usual} \subset W_{h,usual}$. This is *not* true for the spaces of Definition 4.1.

Our spaces can be viewed as generalizations of the Raviart–Thomas spaces RT_{k-1} [29, 31]. Our analysis will make strong use of the RT_{k-1}-projection operator $\tilde{\pi}^{k-1}$. We briefly review RT_{k-1}. By definition,

$$RT_{k-1} = \tilde{\mathbf{V}}_h^{k-1} \times W_h^{k-1},$$

where

$$\tilde{\mathbf{V}}_h^{k-1}(R) = Q_{k,k-1}(R) \times Q_{k-1,k}(R),$$
$$\tilde{\mathbf{V}}_h^{k-1} = \{\mathbf{v} \in H(\text{div}; \Omega) : \mathbf{v}|_R \in \tilde{\mathbf{V}}_h^{k-1}(R) \quad \forall R \in \mathcal{T}_h\}.$$

The operator

$$\tilde{\pi}^{k-1} : (H^1(\Omega))^2 \to \tilde{\mathbf{V}}_h^{k-1}$$

is defined on a rectangle R by the degrees of freedom
  (i) $\langle \tilde{\pi}^{k-1} \mathbf{u} \cdot \nu, \lambda\rangle_e = \langle \mathbf{u} \cdot \nu, \lambda\rangle_e$ for all $\lambda \in P_{k-1}(e)$ and edges $e \subset \partial R$,
  (ii) $(\tilde{\pi}^{k-1} \mathbf{u}, \mathbf{v})_R = (\mathbf{u}, \mathbf{v})_R$ for all $\mathbf{v} \in Q_{k-2,k-1}(R) \times Q_{k-1,k-2}(R)$.

We recall that

$$(5.11) \qquad \nabla \cdot \tilde{\pi}^{k-1} \mathbf{u} = \mathcal{P}_W^{k-1} \nabla \cdot \mathbf{u},$$

$$(5.12) \qquad \|\mathbf{u} - \tilde{\pi}^{k-1} \mathbf{u}\|_0 \le C\|\mathbf{u}\|_r h^r, \quad 1 \le r \le k.$$

Let $\mathcal{P}_{\tilde{V}}^{k-1} : (L^2(\Omega))^2 \to \tilde{\mathbf{V}}_h^{k-1}$ denote projection $\mathcal{P}_{\tilde{V}_h^{k-1}}$. Relations between π^k and $\tilde{\pi}^{k-1}$ are given as follows.

LEMMA 5.3. *Assume that* $\mathbf{u} \in (H^1(\Omega))^2$.
   (a) $\tilde{\pi}^{k-1} \pi^k \mathbf{u} = \tilde{\pi}^{k-1} \mathbf{u}$.
   (b) $\tilde{\pi}^{k-1} \mathbf{u} = \mathcal{P}_{\tilde{V}}^{k-1} \pi^k \mathbf{u}$.
   (c) $\tilde{\pi}^{k-1} \mathbf{u} = \mathcal{P}_{\tilde{V}}^{k-1} \mathbf{u}$, *provided* $\mathbf{u} \in \mathbf{V}_h^k$.

Proof. For (a), π^k preserves the degrees of freedom of $\tilde{\pi}^{k-1}$. Result (b) is a corollary of (a) and (c).

For (c), since $\tilde{\pi}^{k-1}$ is constructed locally on each element, we restrict our analysis to a rectangle R. Assume without loss of generality that $R = \hat{R}$. Since x and y components are independent, we consider only x-components. Then

$$\tilde{V}_{h,1}^{k-1} = \operatorname{span}\{\lambda_i(x)\lambda_j(y) : i \le k \ , \ j \le k-1\},$$
$$V_{h,1}^k = \operatorname{span}\{\lambda_i(x)\lambda_j(y) : i \le k \ , \ j \le k+1\}.$$

We wish to show that for $\mathbf{u} \in \mathbf{V}_h^k(R)$,

$$((\tilde{\pi}^{k-1}\mathbf{u})_1 - u_1, v_1)_R = 0 \quad \forall\, v_1 \in \tilde{V}_{h,1}^{k-1}(R).$$

If $\mathbf{u} \in \tilde{\mathbf{V}}_h^{k-1}(R)$, the result is trivial, so assume

$$u_1 \in \operatorname{span}\{\lambda_i(x)\lambda_j(y) : i \le k, \quad j = k, k+1\}.$$

Then $(u_1, v_1)_R = 0$ for all $v_1 \in \tilde{V}_{h,1}^{k-1}(R)$. Moreover, if e is an edge of ∂R with x constant (i.e., with $\nu = \mathbf{e}_1$), then by the orthogonality of the Legendre polynomials,
   (i) $\langle u_1(x, \cdot), q \rangle_e = 0$ for all $q(y) \in P_{k-1}(e)$,
   (ii) $(u_1, v_1)_R = 0$ for all $v_1 \in Q_{k-2, k-1}(R)$.
These are the degrees of freedom that define the x-component of $\tilde{\pi}^{k-1}$, and so $(\tilde{\pi}^{k-1}\mathbf{u})_1 = 0 = (\mathcal{P}_{\tilde{V}}^{k-1}\mathbf{u})_1$. $\quad\square$

6. A mixed finite element method. To approximate (1.1)–(1.2), (1.4)–(1.5), we first rewrite the equations in mixed variational form. Let

$$\mathbf{V}_0 = \{\mathbf{v} \in H(\operatorname{div}; \Omega) : \mathbf{v} \cdot \nu = 0 \text{ on } \Gamma_N\},$$

and $W = L^2(\Omega)$, unless $a \equiv 0$ and $\Gamma_N = \partial\Omega$ (the pure Neumann problem), in which case $W = \{w \in L^2(\Omega) : \int_\Omega w = 0\}$. In the latter case, we also assume the usual compatibility condition between b and g_N.

We find $(\mathbf{u}, p) \in H(\operatorname{div}; \Omega) \times W$ such that $\mathbf{u} \cdot \nu = g_N$ on Γ_N and

$$(6.1) \qquad (ap, w) + (\nabla \cdot \mathbf{u}, w) = (b, w) \qquad\qquad \forall w \in W,$$

$$(6.2) \qquad (d\mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) = -\langle p_D, \mathbf{v} \cdot \nu \rangle_{\Gamma_D} + (\mathbf{c}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{V}_0.$$

Normally one merely restricts to the finite element spaces to define the mixed method. For our spaces, however, this is not the proper definition of the method

when nonhomogeneous boundary terms appear. We assume that the grid is such that for any edge $e \subset R \in \mathcal{T}_h$, either $e \subset \Gamma_N$ or $e \subset \Gamma_D$.

We begin with Dirichlet conditions. Let us define for edge $e \subset \Gamma_N \cap \partial R$, $R \in \mathcal{T}_h$,

$$\Lambda_h^k(e) = \mathbf{V}_h^k(R) \cdot \nu = P_{k+1}(e),$$
$$\tilde{\Lambda}_h^{k-1}(e) = \tilde{\mathbf{V}}_h^{k-1}(R) \cdot \nu = P_{k-1}(e),$$

and the full spaces Λ_h^k of continuous functions and $\tilde{\Lambda}_h^{k-1}$ of discontinuous functions over Γ_N. With the definition $\mathcal{P}_{\tilde{\Lambda}}^{k-1} = \mathcal{P}_{\tilde{\Lambda}_h^{k-1}}$, we replace p_D by $\mathcal{P}_{\tilde{\Lambda}}^{k-1}p_D$ in the equations.

For Neumann conditions, we need to set $\mathbf{u}_h \cdot \nu$ on Γ_N. This can be done in any reasonable way, but for the error analysis to follow it is convenient to set $\mathbf{u}_h \cdot \nu$ to $\pi^k \mathbf{u}$. Since $\mathbf{u}$ is unknown, we need to define it using only g_N. Near the boundary only, we use the Scott–Zhang [33] modification of the Clément operator $\mathcal{I}^k$ [19] considered in section 5, which has similar properties, except that in (5.1) we must have $r \geq 1$. We set the corner values to a local L^2-projection defined entirely on the boundary of the domain, instead of over rectangles in the domain, so that $\mathcal{I}^k \mathbf{u} \cdot \nu|_{\Gamma_N}$ is defined entirely by $\mathbf{u} \cdot \nu|_{\Gamma_N} = g_N$. Further, we can define an operator $\mathcal{I}_\Lambda^k : L^2(\Gamma_N) \to \Lambda_h^k$ by

$$\mathcal{I}_\Lambda^k g(P) = \mathcal{I}^k g(P)$$

for each corner point P of the grid restricted to Γ_N, and, on each grid edge $e \subset \Gamma_N$,

$$\langle \mathcal{I}_\Lambda^k g, \lambda \rangle_e = \langle g, \lambda \rangle_e \quad \forall \lambda \in P_k(e).$$

Then it is easy to check that

$$\pi^k \mathbf{u} \cdot \nu|_{\Gamma_N} = \mathcal{I}_\Lambda^k g_N.$$

Finally, to set the Neumann boundary condition, we set $\mathbf{u}_h \cdot \nu = \mathcal{I}_\Lambda^k g_N$ on Γ_N. Let

$$\mathbf{V}_{h,0}^k = \{\mathbf{v} \in \mathbf{V}_h^k : \mathbf{v} \cdot \nu = 0 \text{ on } \Gamma_N\}.$$

We now define our mixed finite element method for (6.1)–(6.2). Find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h^k \times W_h^{k-1}$ such that $\mathbf{u}_h \cdot \nu = \mathcal{I}_\Lambda^k g_N$ on Γ_N and

(6.3) $(ap_h, w_h) + (\nabla \cdot \mathbf{u}_h, w_h) = (b, w_h)$ $\hspace{2cm}$ $\forall w_h \in W_h,$

(6.4) $(d\mathbf{u}_h, \mathbf{v}_h) - (p_h, \nabla \cdot \mathbf{v}_h) = -\langle \mathcal{P}_{\tilde{\Lambda}}^{k-1} p_D, \mathbf{v}_h \cdot \nu \rangle_{\Gamma_D} + (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} \mathbf{c}, \mathbf{v}_h)$ $\hspace{0.3cm}$ $\forall \mathbf{v}_h \in \mathbf{V}_{h,0}^k.$

7. Mixed method error analysis. We now analyze the error in approximating (6.1)–(6.2) by (6.3)–(6.4). In (6.2) let $\mathbf{v}$ be replaced by $\tilde{\pi}^{k-1} \mathbf{v}_h$ for $\mathbf{v}_h \in \mathbf{V}_{h,0}^k$. By Lemma 5.3, $\tilde{\pi}^{k-1} \mathbf{v}_h = \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} \mathbf{v}_h$, and with (5.11) we obtain

$$(\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u}, \mathbf{v}_h) - (\mathcal{P}_W^{k-1} p, \nabla \cdot \mathbf{v}_h) = -\langle \mathcal{P}_{\tilde{\Lambda}}^{k-1} p_D, \mathbf{v}_h \cdot \nu \rangle_{\Gamma_D}$$
(7.1) $$+ (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} \mathbf{c}, \mathbf{v}_h), \quad \mathbf{v}_h \in \mathbf{V}_{h,0}^k.$$

Note that in (6.1), if $w \in W_h^{k-1}$, we can replace $\mathbf{u}$ by $\pi^k \mathbf{u}$. Thus the difference of (6.1), (7.1) and (6.3)–(6.4) is

(7.2) $\hspace{0.5cm}$ $(a(p - p_h), w_h) + (\nabla \cdot (\pi^k \mathbf{u} - \mathbf{u}_h), w_h) = 0,$ $\hspace{1cm}$ $w_h \in W_h^{k-1},$

(7.3) $\hspace{0.5cm}$ $(\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}_h, \mathbf{v}_h) - (\mathcal{P}_W^{k-1} p - p_h, \nabla \cdot \mathbf{v}_h) = 0,$ $\hspace{0.3cm}$ $\mathbf{v}_h \in \mathbf{V}_{h,0}^k.$

Select the standard test functions

$$\mathbf{v}_h = \pi^k \mathbf{u} - \mathbf{u}_h \in \mathbf{V}_{h,0}^k \quad \text{and} \quad w_h = \mathcal{P}_W^{k-1} p - p_h \in W_h^{k-1} .$$

Then the sum of (7.2)–(7.3) is

$$(a(p - p_h), p - p_h) + (d(\mathbf{u} - \mathbf{u}_h), \mathbf{u} - \mathbf{u}_h)$$
$$= (a(p - p_h), p - \mathcal{P}_W^{k-1} p) + (d(\mathbf{u} - \mathbf{u}_h), \mathbf{u} - \pi^k \mathbf{u}) + (d\mathbf{u} - \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u}, \pi^k \mathbf{u} - \mathbf{u}_h),$$

from which we conclude, using standard approximation theory [10, 18] and Lemma 5.2, that

$$\|\sqrt{a}(p - p_h)\|_0 + \|d^{1/2}(\mathbf{u} - \mathbf{u}_h)\|_0$$
$$\leq C\{\|p - \mathcal{P}_W^{k-1} p\|_0 + \|\mathbf{u} - \pi^k \mathbf{u}\|_0 + \|d\mathbf{u} - \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u}\|_0\}$$
(7.4) $$\leq C\{\|p\|_r + \|\mathbf{u}\|_r\} h^r$$

for any $1 \leq r \leq k$.

THEOREM 7.1. *If* $\mathbf{u} \in (H^1(\Omega))^2$, *the coefficients* $a \geq 0$ *and* d *are sufficiently smooth, and* d *is uniformly elliptic, then there is some constant* C, *independent of* h, *such that*

(a) $\|\sqrt{a}(p - p_h)\|_0 + \|\mathbf{u} - \mathbf{u}_h\|_0 \leq C\{\|p\|_r + \|\mathbf{u}\|_r\} h^r$, $\quad 1 \leq r \leq k$,

(b) $\|\mathcal{P}_W^{k-1} p - p_h\|_0 \leq C\{\|p\|_r + \|\mathbf{u}\|_{r+1}\} h^{r+1}$, $\quad\quad 0 \leq r \leq k$,

(c) $\|p - p_h\|_0 \leq C\{\|p\|_r + \|\mathbf{u}\|_r\} h^r$, $\quad\quad\quad\quad\quad 1 \leq r \leq k$,

(d) $\|\mathcal{P}_W^{k-1} \nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|_0 \leq C\{\|p\|_r + \|\mathbf{u}\|_r\} h^r$, $\quad 1 \leq r \leq k$,

(e) $\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|_0 \leq C\{\|p\|_r + \|\mathbf{u}\|_r\} h^{r-1}$, $\quad\quad 1 \leq r \leq k$,

where for (b) *and* (c) *we have assumed that problem* (7.5)–(7.7) *below is 2-regular, i.e., its solution satisfies* (7.8), *and for* (e) *we assume that the finite element partition is quasi-uniform.*

While results (c) and (d) are optimal in the rate of convergence, results (a) and (e) are only suboptimal. Result (b) exhibits a superconvergence phenomenon typical of mixed methods. If $k = 1$, we do not control the full divergence error; however, we at least have stability. See, e.g., [25, 27] for conditions that imply 2-regularity.

Proof. We have shown (a) above. For (b) and (c), let $\psi \in L^2(\Omega)$ be such that $\|\psi\|_0 \leq 1$ and consider the solution $\varphi \in H^2(\Omega)$ to

(7.5) $$a\varphi - \nabla \cdot d^{-1} \nabla \varphi = \psi, \quad \Omega,$$

(7.6) $$d^{-1} \nabla \varphi \cdot \nu = 0, \quad \Gamma_N,$$

(7.7) $$\varphi = 0 , \quad \Gamma_D.$$

By hypothesis the problem is 2-regular, so by definition there is some constant $C > 0$ such that

(7.8) $$\|\varphi\|_2 \leq C\|\psi\|_0 \leq C.$$

Then, using Lemma 5.2(d),

$$(\mathcal{P}_W^{k-1} p - p_h, \psi) = (a(\mathcal{P}_W^{k-1} p - p_h), \varphi) - (\mathcal{P}_W^{k-1} p - p_h, \nabla \cdot d^{-1} \nabla \varphi)$$
$$= (a(\mathcal{P}_W^{k-1} p - p_h), \varphi) - (\mathcal{P}_W^{k-1} p - p_h, \nabla \cdot \pi^k d^{-1} \nabla \varphi)$$
$$= (a(\mathcal{P}_W^{k-1} p - p_h), \varphi) - (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}_h, \pi^k d^{-1} \nabla \varphi),$$

by (7.3), since $\pi^k d^{-1} \nabla \varphi \in \mathbf{V}_{h,0}^k$. Our immediate goal is to cancel the coefficients d and d^{-1} in the last term, so we compute

$$
\begin{aligned}
(\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}_h, \pi^k d^{-1}\nabla\varphi) &= (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}_h, \pi^k d^{-1}\nabla\varphi - \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d^{-1}\nabla\varphi) \\
&\quad + (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}_h, \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d^{-1}\nabla\varphi) \\
&= (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}_h, \pi^k d^{-1}\nabla\varphi - \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d^{-1}\nabla\varphi) \\
&\quad + (d(\mathbf{u} - \mathbf{u}_h), \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d^{-1}\nabla\varphi - d^{-1}\nabla\varphi) + (\mathbf{u} - \mathbf{u}_h, \nabla\varphi).
\end{aligned}
$$

We now wish to integrate the last term above by parts and use the other error equation (7.2). However, we must do this carefully, so as to obtain the superconvergence claimed in the theorem. We compute

$$
\begin{aligned}
(\mathbf{u} - \mathbf{u}_h, \nabla\varphi) &= (\mathbf{u} - \pi^k\mathbf{u}, \nabla\varphi) + (\pi^k\mathbf{u} - \mathbf{u}_h, \nabla\varphi - \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1}\nabla\varphi) \\
&\quad + (\pi^k\mathbf{u} - \mathbf{u}_h, \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1}\nabla\varphi).
\end{aligned}
$$

Now, using Lemma 5.3, (5.11), and (7.2),

$$
\begin{aligned}
(\pi^k\mathbf{u} - \mathbf{u}_h, \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1}\nabla\varphi) &= (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1}(\pi^k\mathbf{u} - \mathbf{u}_h), \nabla\varphi) \\
&= (\tilde{\pi}^{k-1}(\pi^k\mathbf{u} - \mathbf{u}_h), \nabla\varphi) \\
&= -(\nabla \cdot \tilde{\pi}^{k-1}(\pi^k\mathbf{u} - \mathbf{u}_h), \varphi) \\
&= -(\nabla \cdot (\pi^k\mathbf{u} - \mathbf{u}_h), \mathcal{P}_W^{k-1}\varphi) \\
&= (a(p - p_h), \mathcal{P}_W^{k-1}\varphi) \\
&= (a(\mathcal{P}_W^{k-1}p - p_h), \mathcal{P}_W^{k-1}\varphi) + ((a - \mathcal{P}_W^0 a)(p - \mathcal{P}_W^{k-1}p), \mathcal{P}_W^{k-1}\varphi),
\end{aligned}
$$

since $\mathcal{P}_W^{k-1}(\mathcal{P}_W^0 a(p - \mathcal{P}_W^{k-1}p)) = 0$. Combining and using the approximation properties of the various projections and (7.4), we obtain that

$$
\begin{aligned}
(\mathcal{P}_W^{k-1}p - p_h, \psi) &= (a(\mathcal{P}_W^{k-1}p - p_h), \varphi - \mathcal{P}_W^{k-1}\varphi) \\
&\quad - (\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}_h, \pi^k d^{-1}\nabla\varphi - \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d^{-1}\nabla\varphi) \\
&\quad - (d(\mathbf{u} - \mathbf{u}_h), \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d^{-1}\nabla\varphi - d^{-1}\nabla\varphi) - (\mathbf{u} - \pi^k\mathbf{u}, \nabla\varphi) \\
&\quad - (\pi^k\mathbf{u} - \mathbf{u}_h, \nabla\varphi - \mathcal{P}_{\tilde{\mathbf{V}}}^{k-1}\nabla\varphi) - ((a - \mathcal{P}_W^0 a)(p - \mathcal{P}_W^{k-1}p), \mathcal{P}_W^{k-1}\varphi) \\
&\leq C \big\{ \big(\|\sqrt{a}(\mathcal{P}_W^{k-1}p - p_h)\|_0 + \|\mathbf{u} - \mathbf{u}_h\|_0 + \|\mathcal{P}_{\tilde{\mathbf{V}}}^{k-1} d\mathbf{u} - d\mathbf{u}\|_0 \\
&\quad + \|\pi^k\mathbf{u} - \mathbf{u}\|_0 + \|a\|_{W^{1,\infty}}\|p - \mathcal{P}_W^{k-1}p\|_0 \big) h + \|\mathbf{u} - \pi^k\mathbf{u}\|_0 \big\} \|\varphi\|_2 \\
&\leq C \big\{ \|p\|_r + \|\mathbf{u}\|_{r+1} \big\} h^{r+1},
\end{aligned}
$$

where $0 \leq r \leq k$. Thus (b) and (c) follow.

Substituting $w_h = \mathcal{P}_W^{k-1}\nabla \cdot (\mathbf{u} - \mathbf{u}_h) \in W_h^{k-1}$ into (7.2) and using (7.4) leads to (d). For (e), we have a standard inverse inequality argument [10]. On a quasi-uniform partition (i.e., one that is quasi-regular and has the size of the largest to smallest rectangle bounded independently of h), we compute

$$
\begin{aligned}
\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|_0 &\leq \|\nabla \cdot (\mathbf{u} - \pi^k\mathbf{u})\|_0 + \|\nabla \cdot (\pi^k\mathbf{u} - \mathbf{u}_h)\|_0 \\
&\leq \|\mathbf{u} - \pi^k\mathbf{u}\|_1 + Ch^{-1}\|\pi^k\mathbf{u} - \mathbf{u}_h\|_0,
\end{aligned}
$$

and the result follows easily.   □

8. The inf-sup condition. Our spaces satisfy the celebrated LBB or inf-sup condition of Ladyzhenskaya [28], Babuška [6], Brezzi [11], and Brezzi and Fortin [15].

THEOREM 8.1. *If Ω is a polygonal domain, then there exists a constant $\gamma > 0$ such that*

$$\inf_{w \in W_h^{k-1}} \sup_{\mathbf{v} \in \mathbf{V}_{h,0}^k} \frac{(\nabla \cdot \mathbf{v}, w)}{\|\mathbf{v}\|\|w\|_0} \geq \gamma > 0,$$

wherein functions of W_h^{k-1} have mean zero if $\Gamma_N = \partial\Omega$ and we take the norm on $\mathbf{V}_h^k$ to be any one of

$$\|\mathbf{v}\| = \left\{\|\mathbf{v}\|_0^2 + \|\nabla \cdot \mathbf{v}\|_0^2\right\}^{1/2} \equiv \|\mathbf{v}\|_{H(\mathrm{div})},$$
$$\|\mathbf{v}\| = \left\{\|\mathbf{v}\|_0^2 + \|\mathcal{P}_W^{k-1}\nabla \cdot \mathbf{v}\|_0^2\right\}^{1/2} \equiv |||v|||,$$

or $\|\mathbf{v}\| = \|\mathbf{v}\|_1$.

Proof. It suffices to show the result for the $(H^1(\Omega))^2$-norm on V_h^k, since

$$|||\mathbf{v}||| \leq \|\mathbf{v}\|_{H(\mathrm{div})} \leq \|\mathbf{v}\|_1.$$

There exists a $\beta > 0$ such that, given $w \in L^2(\Omega)$ (or $w \in L^2(\Omega)/\mathbb{R}$ if $\Gamma_N = \partial\Omega$), there is some $\psi \in (H^1(\Omega))^2$ such that both $\nabla \cdot \psi = w$ and $\|\psi\|_1 \leq \beta\|w\|_0$ (see [5]). Then for $w \in W_h^{k-1}$,

$$\sup_{\mathbf{v} \in \mathbf{V}_{h,0}^k} \frac{(\nabla \cdot \mathbf{v}, w)}{\|\mathbf{v}\|_1\|w\|_0} \geq \frac{(\nabla \cdot \pi^k\psi, w)}{\|\pi^k\psi\|_1\|w\|_0} \geq \frac{\|w\|_0}{C\|\psi\|_1} \geq \frac{1}{\beta C} > 0,$$

since π^k is bounded on $(H^1(\Omega))^2$. $\square$

We analyzed our finite element method in a direct way in section 7 above, rather than use the inf-sup theory of saddle point problems [6, 10, 11, 15]. This is because it is not straightforward to apply the theory. To do so would require both that the form $(\nabla \cdot \mathbf{v}, w)$ be continuous and that $(d\mathbf{v}, \mathbf{v})$ be coercive on the set $Z = \{\mathbf{v} \in \mathbf{V}_{h,0}^k : (\nabla \cdot \mathbf{v}, w) = 0 \ \forall w \in W_h^{k-1}\}$. For the former, we would need to take the norm on $\mathbf{V}_h^k$ to be at least $\|\cdot\|_{H(\mathrm{div})}$, while for the latter we would need the norm to be no stronger than $|||\cdot|||$. The problem is that testing $(\nabla \cdot \mathbf{v}, w)$ by $w \in W_h^{k-1}$ does not control the full divergence of $\mathbf{v}$, but only its projection $\mathcal{P}_W^{k-1}\nabla \cdot \mathbf{v}$.

9. Numerical examples. In this section we present some numerical results on the problem (1.1)–(1.2), (1.5) with $\Omega = (0,1)^2$, $\Gamma_D = \partial\Omega$, $a = 0$, and $\mathbf{c} = 0$. We fix the true solution and then define p_D and b so that the equations are satisfied. The test cases are summarized in Table 9.1, wherein d and p are defined, as well as a statement as to whether the grid is uniform or not. The observed convergence errors are shown in Table 9.2, and in Table 9.3 we show the observed convergence rates. The norms of the errors were computed using a tensor product 3-point Gauss rule, and the convergence rates were obtained by fitting the norms of the errors to $\exp(m \log h + b)$, with m being the convergence rate.

As can be seen from Table 9.3, p converges to p_h as $O(h)$, and $\mathcal{P}_W^0 p$ as $O(h^2)$, both as expected from Theorem 7.1. We also see that $\mathbf{u}$ converges to $\mathbf{u}_h$ somewhat better than expected. It appears that on the uniform grid we attain $O(h^{3/2})$ super-convergence in the L^2-norm, and $O(h^{1/2})$ in the H^1-seminorm. On the random grid, the superconvergence appears to be lost, at least mostly so. For these test cases,

TABLE 9.1
Numerical test cases.

Case	Coefficient d	True solution	Grid
1	1	$y^4 e^x$	uniform
2	1	$\dfrac{\cos(x^2 y)}{x^2 + xy + 1}$	uniform
3	$\begin{pmatrix} e^{2xy^2} & 0 \\ 0 & 1/(1.1 + x^2 - y) \end{pmatrix}$	$\dfrac{\cos(x^2 y)}{x^2 + xy + 1}$	uniform
4	$\begin{pmatrix} e^{2xy^2} & 0 \\ 0 & 1/(1.1 + x^2 - y) \end{pmatrix}$	$\dfrac{\cos(x^2 y)}{x^2 + xy + 1}$	random

TABLE 9.2
Observed errors.

Case	h	$\|p - p_h\|_0$	$\|\mathcal{P}_W^0 p - p_h\|_0$	$\|\mathbf{u} - \mathbf{u}_h\|_0$	$\|\nabla(\mathbf{u} - \mathbf{u}_h)\|_0$	$\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|_0$
1	1/8	9.90e−2	1.86e−3	8.11e−2	3.43e+0	9.40e−1
	1/16	4.98e−2	4.31e−4	3.01e−2	2.60e+0	4.60e−1
	1/32	2.49e−2	1.03e−4	1.09e−2	1.92e+0	2.27e−1
	1/64	1.25e−2	2.64e−5	3.92e−3	1.38e+0	1.13e−1
2	1/8	2.55e−2	2.57e−4	1.23e−2	5.16e−1	9.82e−2
	1/16	1.27e−2	5.95e−5	4.51e−3	3.91e−1	4.80e−2
	1/32	6.37e−3	1.43e−5	1.62e−3	2.86e−1	2.38e−2
	1/64	3.18e−3	3.22e−6	5.78e−4	2.05e−1	1.18e−2
3	1/8	2.55e−2	4.57e−4	3.06e−2	1.26e+0	2.88e−1
	1/16	1.27e−2	1.17e−4	1.17e−2	9.93e−1	1.57e−1
	1/32	6.37e−3	2.96e−5	4.29e−3	7.45e−1	8.00e−2
	1/64	3.18e−3	6.30e−6	1.54e−3	5.43e−1	4.00e−2
4	1/8	2.68e−2	4.81e−4	3.17e−2	1.23e+0	3.00e−1
	1/16	1.34e−2	1.18e−4	1.20e−2	9.95e−1	1.50e−1
	1/32	6.68e−3	3.38e−5	5.37e−3	8.73e−1	8.03e−2
	1/64	3.38e−3	5.80e−6	2.44e−3	6.85e−1	4.04e−2

TABLE 9.3
Observed convergence rates. Best fit to $y = \exp(m \log h + b)$, with m reported below (where y is the norm of the error).

Case	$\|p - p_h\|_0$	$\|\mathcal{P}_W^0 p - p_h\|_0$	$\|\mathbf{u} - \mathbf{u}_h\|_0$	$\|\nabla(\mathbf{u} - \mathbf{u}_h)\|_0$	$\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|_0$
1	0.996	2.049	1.457	0.437	1.020
2	1.000	2.100	1.471	0.445	1.018
3	1.000	2.052	1.438	0.405	0.950
4	0.997	2.092	1.227	0.271	0.958

$\mathcal{P}_W^0 \nabla \cdot \mathbf{u}_h = \mathcal{P}_W^0 \nabla \cdot \mathbf{u}$. However, it appears that $\nabla \cdot \mathbf{u}$ approaches $\nabla \cdot \mathbf{u}_h$ with rate $O(h)$. We have not been able to demonstrate that this is true in general.

If the projection of the Dirichlet data is *not* used in (6.4), then the results degrade significantly. Thus, it is important to include this projection in the method. In our tests, it amounts to simplifying the computation by replacing the Dirichlet value on each boundary element edge by its average value.

10. The spaces in three dimensions. There are several ways to define our mixed finite element spaces in three dimensions. We present the version with as few degrees of freedom as seems possible.

Let

$$\tilde{Q}_{j,j} = \{p \in Q_{j,j} : p \text{ has exact degree } j\}.$$

Then $\dim Q_{j,j} = 2j + 1$. Moreover, let

$$\tilde{Q}_{i,j,j}^{x} = \{p \in Q_{i,j,j} : p \text{ has exact degree } j \text{ in } y \text{ and } z\},$$

and define similarly $\tilde{Q}_{j,i,j}^{y}$ and $\tilde{Q}_{j,j,i}^{z}$, each with dimension $(i+1)(2j+1)$.

We next define some spaces of "bubble" functions. For each integer $k \geq 1$, let $R = (a_0, a_1) \times (b_0, b_1) \times (c_0, c_1)$ and

$$
\begin{aligned}
B_x^k(R) = \big\{ p : p(x,y,z) \\
= (y - b_0)(b_1 - y)(z - c_0)(c_1 - z)\big[(a_1 - x)\,q_0(y,z) + (x - a_0)\,q_1(y,z)\big] \\
\text{for some } q_0, q_1 \in \tilde{Q}_{k-1,k-1}((b_0,b_1) \times (c_0,c_1))\big\} \\
\subset Q_{1,k+1,k+1}(R),
\end{aligned}
$$

with a similar definition for $B_y^k(R)$ and $B_z^k(R)$. Also let

$$
\begin{aligned}
\mathbf{B}^k(R) = \big\{\mathbf{v} : \mathbf{v}(x,y,z) = (x - a_0)(a_1 - x)(y - b_0)(b_1 - y)(z - c_0)(c_1 - z)\,\psi(x,y,z) \\
\text{for some } \psi \in \tilde{Q}_{k-2,k-1,k-1}^{x}(R) \times \tilde{Q}_{k-1,k-2,k-1}^{y}(R) \times \tilde{Q}_{k-1,k-1,k-2}^{z}(R)\big\} \\
\subset Q_{k,k+1,k+1}(R) \times Q_{k+1,k,k+1}(R) \times Q_{k+1,k+1,k}(R).
\end{aligned}
$$

DEFINITION 10.1. *For each integer $k \geq 1$, for any $R \in \mathcal{T}_h$, let*

$$
\begin{aligned}
\mathbf{V}_h^k(R) &= \big(Q_{k,k,k}(R)\big)^3 + B_x^k(R) \times B_y^k(R) \times B_z^k(R) + \mathbf{B}^k(R) \\
&\subset Q_{k,k+1,k+1}(R) \times Q_{k+1,k,k+1}(R) \times Q_{k+1,k+1,k}(R), \\
W_h^{k-1}(R) &= Q_{k-1,k-1,k-1}(R),
\end{aligned}
$$

and define $\mathbf{V}_h^k \times W_h^{k-1}$ as in Definition 4.1.

LEMMA 10.2. *Let R be a rectangular parallelepiped in three dimensions. For each edge $e \in \partial R$, fix a set $\mathcal{S}_e$ of $k+1$ distinct points, including the endpoints. Then $\mathbf{u}_h \in \mathbf{V}_h^k(R)$ is uniquely defined by the following degrees of freedom:*

(i) *For each edge $e \in \partial R$, point $P \in \mathcal{S}_e$, and Cartesian direction $j = 1, 2, 3$,*

$$DOF_{P,j}^{(\mathrm{i})}(\mathbf{u}_h) = u_{h,j}(P) = \mathbf{u}_h(P) \cdot \mathbf{e}_j.$$

(ii) *On each face $f \subset \partial R$,*

$$DOF_{f,\tau_f,\lambda}^{(\mathrm{ii})}(\mathbf{u}_h) = \langle \mathbf{u}_h \cdot \tau_f, \lambda \rangle_f \quad \forall \lambda \in Q_{k-2,k-2}(f),$$

where τ_f is one of the two linearly independent Cartesian unit tangential directions.

(iii) *On each face $f \subset \partial R$,*

$$DOF_{f,\lambda}^{(\mathrm{iii})}(\mathbf{u}_h) = \langle \mathbf{u}_h \cdot \nu_f, \lambda \rangle_f \quad \forall \lambda \in Q_{k-1,k-1}(f),$$

where ν_f is the outer unit normal direction.

(iv) *Over R,*

$$
\begin{aligned}
DOF_{\mathbf{v}}^{(\mathrm{iv})}(\mathbf{u}_h) = (\mathbf{u}_h, \mathbf{v})_R \quad \forall \mathbf{v} \in Q_{k-2,k-1,k-1}(R) \times Q_{k-1,k-2,k-1}(R) \\
\times Q_{k-1,k-1,k-2}(R).
\end{aligned}
$$

Moreover, on each edge $e \subset \partial R$, $\mathbf{u}_h|_e$ is uniquely defined by the degrees of freedom (i) *restricted to e, and on each face $f \subset \partial R$, $\mathbf{u}_h|_f$ is uniquely defined by the degrees of freedom* (i)–(iii) *restricted to f.*

The dimension of $\mathbf{V}_h^k(R)$ is

$$
\begin{aligned}
\dim \mathbf{V}_h^k(R) &= 3(k+1)^3 + 6(2k-1) + 3(k-1)(2k-1) = 3(k^3 + 5k^2 + 4k) \\
&= 3k(k+1)(k+4).
\end{aligned}
$$

The number of independent degrees of freedom for (i) is $3(12(k-1)+8) = 36k - 12$, for (ii) is $12(k-1)^2 = 12k^2 - 24k + 12$, for (iii) is $6k^2$, and for (iv) is $3k^2(k-1)$, for a total matching the dimension of the space. Therefore the proof of the lemma is similar to that for Lemma 4.2, and we omit it.

The theoretical results of the previous sections can be developed for these spaces in a relatively straightforward way. In particular, we have analogues of the definition of π^k, Lemmas 5.2 and 5.3, the development of the mixed finite element method, and Theorems 7.1 and 8.1.

REFERENCES

[1] T. ARBOGAST AND D. S. BRUNSON, *A computational method for approximating a Darcy–Stokes system governing a vuggy porous medium*, submitted.

[2] T. ARBOGAST, D. S. BRUNSON, S. L. BRYANT, AND J. J. W. JENNINGS, *A Computational Investigation of a Macro-Model for Vuggy Porous Media*, manuscript.

[3] T. ARBOGAST, D. S. BRUNSON, S. L. BRYANT, AND J. W. JENNINGS, *A preliminary computational investigation of a macro-model for vuggy porous media*, in Computational Methods in Water Resources XV, C. T. Miller, M. W. Farthing, W. G. Gray, and G. F. Pindar, eds., Elsevier, New York, 2004.

[4] T. ARBOGAST AND M. F. WHEELER, *A characteristics-mixed finite element method for advection-dominated transport problems*, SIAM J. Numer. Anal., 32 (1995), pp. 404–424.

[5] D. N. ARNOLD, L. R. SCOTT, AND M. VOGELIUS, *Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon*, Ann. Scuola Norm. Sup. Pisa Cl. Sci.-Serie IV, 15 (1988), pp. 169–192.

[6] I. BABUŠKA, *The finite element method with Lagrangian multipliers*, Numer. Math., 20 (1973), pp. 179–192.

[7] J. BEAR, *Dynamics of Fluids in Porous Media*, Dover, New York, 1972.

[8] G. S. BEAVERS AND D. D. JOSEPH, *Boundary conditions at a naturally permeable wall*, J. Fluid Mech., 30 (1967), pp. 197–207.

[9] C. BERNARDI AND G. RAUGEL, *Analysis of some finite elements for the Stokes problem*, Math. Comp., 44 (1985), pp. 71–79.

[10] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[11] F. BREZZI, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, RAIRO Oper. Res., 8 (1974), pp. 129–151.

[12] F. BREZZI, J. DOUGLAS, JR., R. DURÀN, AND M. FORTIN, *Mixed finite elements for second order elliptic problems in three variables*, Numer. Math., 51 (1987), pp. 237–250.

[13] F. BREZZI, J. DOUGLAS, JR., M. FORTIN, AND L. D. MARINI, *Efficient rectangular mixed finite elements in two and three space variables*, RAIRO Modél. Math. Anal. Numér., 21 (1987), pp. 581–604.

[14] F. BREZZI, J. DOUGLAS, JR., AND L. D. MARINI, *Two families of mixed elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235.

[15] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[16] F. BREZZI, M. FORTIN, AND L. D. MARINI, *Mixed finite element methods with continuous stresses*, Math. Models Methods Appl. Sci., 3 (1993), pp. 275–287.

[17] Z. CHEN AND J. DOUGLAS, JR., *Prismatic mixed finite elements for second order elliptic problems*, Calcolo, 26 (1989), pp. 135–148.

[18] P. CIARLET, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.

[19] P. CLÉMENT, *Approximation by finite element functions using local regularization*, RAIRO Anal. Numèr., 9 (1975), pp. 77–84.

[20] C. N. DAWSON, T. F. RUSSELL, AND M. F. WHEELER, *Some improved error estimates for the modified method of characteristics*, SIAM J. Numer. Anal., 26 (1989), pp. 1487–1512.

[21] J. DOUGLAS, JR., AND J. E. ROBERTS, *Global estimates for mixed methods for second order elliptic equations*, Math. Comp., 44 (1985), pp. 39–52.

[22] J. DOUGLAS, JR., AND T. F. RUSSELL, *Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures*, SIAM J. Numer. Anal., 19 (1982), pp. 871–885.

[23] R. E. EWING, T. F. RUSSELL, AND M. F. WHEELER, *Convergence analysis of an approximation of miscible displacement in porous media by mixed finite elements and a modified method of characteristics*, Comput. Methods Appl. Mech. Engrg., 47 (1984), pp. 73–92.

[24] M. FORTIN, *Old and new finite elements for incompressible flows*, Internat. J. Numer. Methods Fluids, 1 (1981), pp. 347–364.

[25] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 1983.

[26] V. GIRAULT AND P. A. RAVIART, *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, Springer-Verlag, Berlin, 1986.

[27] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.

[28] O. A. LADYZHENSKAYA, *The Mathematical Theory of Viscous Incompressible Flow*, 2nd ed., Gordon and Breach, New York, 1969.

[29] J. C. NÉDÉLEC, *Mixed finite elements in* $\mathbf{R}^3$, Numer. Math., 35 (1980), pp. 315–341.

[30] D. W. PEACEMAN, *Fundamentals of Numerical Reservoir Simulation*, Elsevier, Amsterdam, 1977.

[31] R. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for* 2*nd order elliptic problems*, in Mathematical Aspects of Finite Element Methods, I. Galligani and E. Magenes, eds., Lecture Notes in Math. 606, Springer-Verlag, New York, 1977, pp. 292–315.

[32] T. F. RUSSELL AND M. F. WHEELER, *Finite element and finite difference methods for continuous flows in porous media*, in The Mathematics of Reservoir Simulation, R. E. Ewing, ed., Frontiers Appl. Math. 1, SIAM, Philadelphia, 1983, pp. 35–106.

[33] L. R. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[34] J. M. THOMAS, *Sur l'analyse numerique des methodes d'elements finis hybrides et mixtes*, Ph.D. thesis, Sciences Mathematiques, à l'Universite Pierre et Marie Curie, Paris, 1977.