QUADRILATERAL H(div) FINITE ELEMENTS[*]

DOUGLAS N. ARNOLD[†], DANIELE BOFFI[‡], AND RICHARD S. FALK[§]

Abstract. We consider the approximation properties of quadrilateral finite element spaces of vector fields defined by the Piola transform, extending results previously obtained for scalar approximation. The finite element spaces are constructed starting with a given finite dimensional space of vector fields on a square reference element, which is then transformed to a space of vector fields on each convex quadrilateral element via the Piola transform associated to a bilinear isomorphism of the square onto the element. For affine isomorphisms, a necessary and sufficient condition for approximation of order $r + 1$ in L^2 is that each component of the given space of functions on the reference element contain all polynomial functions of total degree at most r. In the case of bilinear isomorphisms, the situation is more complicated and we give a precise characterization of what is needed for optimal order L^2-approximation of the function and of its divergence. As applications, we demonstrate degradation of the convergence order on quadrilateral meshes as compared to rectangular meshes for some standard finite element approximations of $\boldsymbol{H}$(div). We also derive new estimates for approximation by quadrilateral Raviart–Thomas elements (requiring less regularity) and propose a new quadrilateral finite element space which provides optimal order approximation in $\boldsymbol{H}$(div). Finally, we demonstrate the theory with numerical computations of mixed and least squares finite element approximations of the solution of Poisson's equation.

Key words. quadrilateral, finite element, approximation, mixed finite element

AMS subject classifications. 65N30, 41A10, 41A25, 41A27, 41A63

DOI. 10.1137/S0036142903431924

1. Introduction. Many mixed finite element methods are based on variational principles employing the space $\boldsymbol{H}(\mathrm{div}, \Omega)$ consisting of $\boldsymbol{L}^2$ vector fields with divergence in L^2. For such methods, finite element subspaces of $\boldsymbol{H}(\mathrm{div}, \Omega)$ are generally constructed starting from a space of reference shape functions on a reference element, typically the unit simplex or unit square in two dimensions. See, e.g., [4] for numerous examples. These shape functions are then transformed to general triangular, rectangular, or quadrilateral elements via polynomial diffeomorphisms and the Piola transform. For the case of triangular and rectangular (or more generally parallelogram) elements, i.e., the case of affine isomorphisms, the order of approximation so achieved can be easily determined from the highest degree of complete polynomial space contained in the space of reference shape functions. In the case of arbitrary convex quadrilaterals with bilinear diffeomorphisms, the situation is less well understood. In this paper, we determine precisely what reference shape functions are needed to obtain a given order of approximation in $\boldsymbol{L}^2$ and $\boldsymbol{H}(\mathrm{div}, \Omega)$ by such elements. It turns out that the accuracy of some of the standard $\boldsymbol{H}(\mathrm{div}, \Omega)$ finite elements is lower for general quadrilateral elements than for rectangular elements.

Let $\hat{K}$ be a reference element, the closure of an open set in $\mathbb{R}^2$, and let $\boldsymbol{F} : \hat{K} \to \mathbb{R}^2$ be a diffeomorphism of $\hat{K}$ onto an actual element $K = \boldsymbol{F}(\hat{K})$. For functions in

$\boldsymbol{H}(\mathrm{div}, \Omega)$ the natural way to transform functions from $\hat{K}$ to K is via the *Piola transform*. Namely, given a function $\hat{\boldsymbol{u}} : \hat{K} \to \mathbb{R}^2$, we define $\boldsymbol{u} = \boldsymbol{P}_F \hat{u} : K \to \mathbb{R}^2$ by

$$(1.1) \qquad \boldsymbol{u}(\boldsymbol{x}) = J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1} D\boldsymbol{F}(\hat{\boldsymbol{x}}) \hat{\boldsymbol{u}}(\hat{\boldsymbol{x}}),$$

where $\boldsymbol{x} = \boldsymbol{F}(\hat{\boldsymbol{x}})$, and $D\boldsymbol{F}(\hat{\boldsymbol{x}})$ is the Jacobian matrix of the mapping $\boldsymbol{F}$ and $J\boldsymbol{F}(\hat{\boldsymbol{x}})$ its determinant. The transform has the property that if $\boldsymbol{u} = \boldsymbol{P}_F \hat{\boldsymbol{u}}$, $p = \hat{p} \circ \boldsymbol{F}^{-1}$ for some $\hat{p} : \hat{K} \to \mathbb{R}$, and $\boldsymbol{n}$ and $\hat{\boldsymbol{n}}$ denote the unit outward normals on ∂K and $\partial \hat{K}$, respectively, then

$$\int_K \mathrm{div}\,\boldsymbol{u} p \, d\boldsymbol{x} = \int_{\hat{K}} \hat{\mathrm{div}}\,\hat{\boldsymbol{u}} \hat{p} \, d\hat{\boldsymbol{x}}, \qquad \int_{\partial K} \boldsymbol{u} \cdot \boldsymbol{n} p \, ds = \int_{\partial \hat{K}} \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}} \hat{p} \, d\hat{s}.$$

Since continuity of $\boldsymbol{u} \cdot \boldsymbol{n}$ is necessary for finite element subspaces of $\boldsymbol{H}(\mathrm{div}, \Omega)$, use of the Piola transform facilitates the definition of finite element subspaces of $\boldsymbol{H}(\mathrm{div}, \Omega)$ by mapping from a reference element. Another important property of the Piola transform, which follows directly from the chain rule and which we shall use frequently below, is that if $\boldsymbol{G}$ is a diffeomorphism whose domain is K, then

$$(1.2) \qquad \boldsymbol{P}_{G \circ F} = \boldsymbol{P}_G \circ \boldsymbol{P}_F.$$

Using the Piola transform, a standard construction of a finite element subspace proceeds as follows. Let $\hat{K}$ be a fixed reference element, typically either the unit simplex or the unit square. Let $\hat{\boldsymbol{V}} \subset \boldsymbol{H}(\mathrm{div}, \hat{K})$ be a finite-dimensional space of vector fields on $\hat{K}$, typically polynomial, the space of reference *shape functions*. Now suppose we are given a mesh $\mathcal{T}_h$ consisting of elements K, each of which is the image of $\hat{K}$ under some given diffeomorphism: $K = \boldsymbol{F}_K(\hat{K})$. Via the Piola transform we then obtain the space $\boldsymbol{P}_{F_K} \hat{\boldsymbol{V}}$ of shape functions on K. Finally we define the finite element space as

$$\boldsymbol{S}_h = \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega) \mid \boldsymbol{v}|_K \in \boldsymbol{P}_{F_K} \hat{\boldsymbol{V}} \; \forall K \in \mathcal{T}_h\}.$$

Recall that $\boldsymbol{S}_h$ may be characterized as the subspace of

$$\boldsymbol{V}_h := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v}|_K \in \boldsymbol{P}_{F_K} \hat{\boldsymbol{V}} \; \forall K \in \mathcal{T}_h\},$$

consisting of vector fields whose normal component is continuous across interelement edges.

We now recall a few examples of this construction in the case where $\hat{K}$ is the unit square. If we restrict to linear diffeomorphisms $\boldsymbol{F}$, the resulting finite elements $K = \boldsymbol{F}(\hat{K})$ will be parallelograms (or, with the further restriction to diagonal linear diffeomorphisms, rectangles). If we allow general bilinear diffeomorphisms, the resulting finite elements can be arbitrary convex quadrilaterals. The best known example of shape functions on the reference square for construction of $\boldsymbol{H}(\mathrm{div}, \Omega)$ finite element spaces is the Raviart–Thomas space of index $r \geq 0$ for which $\hat{\boldsymbol{V}}$ is taken to be $\boldsymbol{\mathcal{RT}}_r := \mathcal{P}_{r+1,r}(\hat{K}) \times \mathcal{P}_{r,r+1}(\hat{K})$. Here and below $\mathcal{P}_{s,t}(\hat{K})$ denotes the space of polynomial functions on $\hat{K}$ of degree at most s in $\hat{x}_1$ and at most t in $\hat{x}_2$. Thus a basis for $\boldsymbol{\mathcal{RT}}_r$ is given by the $2(r+1)(r+2)$ vector fields

$$(1.3) \qquad (\hat{x}_1^i \hat{x}_2^j, 0), \quad (0, \hat{x}_1^j \hat{x}_2^i), \quad 0 \leq i \leq r+1, \; 0 \leq j \leq r.$$

A second example is given by choosing $\hat{\boldsymbol{V}}$ to be the Brezzi–Douglas–Marini space of index $r \geq 1$, $\hat{\boldsymbol{V}} = \boldsymbol{\mathcal{BDM}}_r$, which is the span of $\boldsymbol{\mathcal{P}}_r(\hat{K})$ and the two additional vector fields $\mathrm{curl}(\hat{x}_1^{r+1}\hat{x}_2)$ and $\mathrm{curl}(\hat{x}_1\hat{x}_2^{r+1})$. Another possibility is the Brezzi–Douglas–Fortin–Marini space $\hat{\boldsymbol{V}} = \boldsymbol{\mathcal{BDFM}}_{r+1}$, $r \geq 0$, which is the subspace of codimension 2 of $\boldsymbol{\mathcal{P}}_{r+1}(\hat{K})$ spanned by $(\hat{x}_1^i\hat{x}_2^j, 0)$ and $(0, \hat{x}_1^j\hat{x}_2^i)$ for nonnegative i and j with $i+j \leq r+1$ and $j \leq r$. We note that for each of these choices $\hat{\boldsymbol{V}}$ strictly contains $\boldsymbol{\mathcal{P}}_r(\hat{K})$ but does not contain $\boldsymbol{\mathcal{P}}_{r+1}(\hat{K})$. Note that $\boldsymbol{\mathcal{BDM}}_0$ is not defined, $\boldsymbol{\mathcal{BDFM}}_1 = \boldsymbol{\mathcal{RT}}_0$, and $\boldsymbol{\mathcal{BDM}}_r \subsetneqq \boldsymbol{\mathcal{BDFM}}_{r+1} \subsetneqq \boldsymbol{\mathcal{RT}}_r$ for $r \geq 1$. More information about these spaces can be found in [4, section III.3.2].

One of the basic issues in finite element theory concerns the approximation properties of finite element spaces. Namely, under certain regularity assumptions on the mesh $\mathcal{T}_h$, for a given smooth vector field $\boldsymbol{u} : \Omega \to \mathbb{R}^2$ one usually estimates the error (in some norm to be made more precise) in the best approximation of $\boldsymbol{u}$ by vector fields in $\boldsymbol{S}_h$ as a quantity involving powers of h, the maximum element diameter. For instance, given a shape-regular sequence of triangular or parallelogram meshes $\mathcal{T}_h$ of Ω with $\boldsymbol{S}_h$ the corresponding Raviart–Thomas spaces of index $r \geq 0$, then for any vector field $\boldsymbol{u}$ smooth enough that the right-hand sides of the next expressions make sense, there exists $\boldsymbol{\pi}_h \boldsymbol{u} \in \boldsymbol{S}_h$ such that (cf. [4])

$$\|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^{r+1} |\boldsymbol{u}|_{H^{r+1}(\Omega)},$$
$$\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u})\|_{L^2(\Omega)} \leq C h^{r+1} |\mathrm{div}\,\boldsymbol{u}|_{H^{r+1}(\Omega)}.$$

In the case of more general shape-regular convex quadrilaterals, the best known estimate appears to be the one obtained by Thomas in [12]:

$$\|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^{r+1} [|\boldsymbol{u}|_{H^{r+1}(\Omega)} + h|\mathrm{div}\,\boldsymbol{u}|_{H^{r+1}(\Omega)}],$$
$$\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u})\|_{L^2(\Omega)} \leq C h^r |\mathrm{div}\,\boldsymbol{u}|_{H^{r+1}(\Omega)}.$$

Note that the order in h for the L^2 estimate on $\boldsymbol{u}$ is the same as for the parallelogram meshes, but additional regularity is required, while the estimate for div $\boldsymbol{u}$ is one order lower in h. As we shall see below, the latter estimate cannot be improved. However, in section 4 of this paper, we use a modification of the usual scaling argument to obtain the improved L^2 estimate

$$\|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^{r+1} |\boldsymbol{u}|_{H^{r+1}(\Omega)}.$$

We restrict our presentation to two-dimensional domains, the three-dimensional case being considerably more complicated. We hope to address this issue in future work. We observe that in [9] the construction of $\boldsymbol{H}(\mathrm{div}, \Omega)$ elements on hexahedrons has been considered. The point of view of [9] is somewhat different from ours in that the elements are not obtained by applying the Piola transform starting from a fixed set of basis functions on the unit cube. Other papers dealing with modifications of standard shape functions for the approximation of vector fields are [11, 8]; in the first paper a simple lowest-order two-dimensional element is proposed (which is not obtained via the Piola transform), while in the second paper a construction based on macroelements is presented.

In this paper, we adapt the theory presented in [1] to the case of vector elements defined by the Piola transform, seeking necessary conditions for L^2-approximation of order $r+1$ for $\boldsymbol{u}$ and div $\boldsymbol{u}$. More specifically, we shall prove in section 3 that in order for the $\boldsymbol{L}^2$ error in the best approximation of $\boldsymbol{u}$ by functions in $\boldsymbol{V}_h$ to be of order

$r + 1$, the space $\hat{\boldsymbol{V}}$ must contain $\boldsymbol{\mathcal{S}}_r$, where $\boldsymbol{\mathcal{S}}_r$ is the subspace of codimension one of $\boldsymbol{\mathcal{RT}}_r$ spanned by the vector fields in (1.3) except that the two fields $(\hat{x}_1^{r+1}\hat{x}_2^r, 0)$ and $(0, \hat{x}_1^r\hat{x}_2^{r+1})$ are replaced by the single vector field $(\hat{x}_1^{r+1}\hat{x}_2^r, -\hat{x}_1^r\hat{x}_2^{r+1})$. To establish this result, we shall exhibit a domain Ω and a sequence $\mathcal{T}_h$ of meshes of it, and prove that whenever $\boldsymbol{\mathcal{S}}_r$ is not contained in $\hat{\boldsymbol{V}}$, there exists a smooth vector field $\boldsymbol{u}$ on Ω such that

$$\inf_{\boldsymbol{v}\in\boldsymbol{V}_h} \|\boldsymbol{u} - \boldsymbol{v}\|_{L^2(\Omega)} \neq o(h^r).$$

The example is far from pathological. The domain is simply a square, the mesh sequence does not degenerate in any sense—in fact all the elements of all the meshes in the sequence are similar to a single right trapezoid—and the function $\boldsymbol{u}$ is a polynomial. We use the same mesh sequence to establish a necessary condition for order $r + 1$ approximation to div $\boldsymbol{u}$, namely that div $\hat{\boldsymbol{V}} \supseteq \mathcal{R}_r$, where $\mathcal{R}_r$ is the subspace of codimension one of $\mathcal{Q}_{r+1}$, the space of polynomials of degree $\leq r + 1$ in each variable separately, spanned by the monomials in $\mathcal{Q}_{r+1}$ except $\hat{x}_1^{r+1}\hat{x}_2^{r+1}$. A consequence of these results, also discussed in section 3, is that while the Raviart–Thomas space of index r achieves order $r + 1$ approximation in $\boldsymbol{L}^2$ for quadrilateral meshes as for rectangular meshes, the order of approximation of the divergence is only of order r in the quadrilateral case (but of order $r + 1$ for rectangular meshes). Thus, in the case $r = 0$, there is no convergence in $\boldsymbol{H}(\mathrm{div}, \Omega)$. For the Brezzi–Douglas–Marini and Brezzi–Douglas–Fortin–Marini spaces, the order of convergence is severely reduced on general quadrilateral meshes not only for div $\boldsymbol{u}$ but also for $\boldsymbol{u}$.

In section 4, we show that the necessary conditions for order $r+1$ approximation of $\boldsymbol{u}$ and div $\boldsymbol{u}$ established in section 3 are also sufficient. The argument used allows us to obtain the previously mentioned improved estimate for approximation by quadrilateral Raviart–Thomas elements. In section 5, we devise a new finite element subspace of $\boldsymbol{H}(\mathrm{div}, \Omega)$ which gives optimal order approximation in both $\ll^2$ and $\boldsymbol{H}(\mathrm{div}, \Omega)$ on general convex quadrilaterals. In sections 6 and 7, we present applications of these results to the approximation of second order elliptic partial differential equations by mixed and least squares finite element methods. In particular, we show that despite the lower order of approximation of the divergence by Raviart–Thomas quadrilateral elements, the mixed method approximation of the scalar and vector variable retain optimal order convergence orders in L^2. By contrast, error estimates for the least squares method indicate a possible loss of convergence for both the scalar and vector variable. In the final section, we illustrate the positive results with some numerical examples and confirm the degradation of accuracy on quadrilateral meshes in the cases predicted by our theory.

2. Approximation theory of vector fields on rectangular meshes. In this preliminary section of the paper we adapt to vector fields the results presented in the corresponding section of [1] for scalar functions. Although the Piola transform is used in the definition of the finite elements, its simple expression on rectangular meshes requires only minor changes in the proof given in [1] and so we give only a statement of the results.

Let K be any square with edges parallel to the axes, namely $K = \boldsymbol{F}_K(\hat{K})$ with

$$(2.1) \qquad\qquad \boldsymbol{F}_K(\hat{\boldsymbol{x}}) = \boldsymbol{x}_K + h_K\hat{\boldsymbol{x}},$$

where $\boldsymbol{x}_K \in \mathbb{R}^2$ is the lower left corner of K and $h_K > 0$ is its side length. The Piola transform of $\hat{\boldsymbol{u}} \in \boldsymbol{L}^2(\hat{K})$ is simply given by $(\boldsymbol{P}_{F_K}\hat{\boldsymbol{u}})(\boldsymbol{x}) = h_K^{-1}\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}})$ where

$\boldsymbol{x} = \boldsymbol{F}_K\hat{\boldsymbol{x}}$. We also have the simple expressions $\mathrm{div}(\boldsymbol{P}_{F_K}\hat{\boldsymbol{u}})(\boldsymbol{x}) = h_K^{-2}\,\hat{\mathrm{div}}\,\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}})$ and $\|\boldsymbol{P}_{F_K}\hat{\boldsymbol{u}}\|_{L^2(K)} = \|\hat{\boldsymbol{u}}\|_{L^2(\hat{K})}$.

Let Ω denote the unit square (Ω and $\hat{K}$ both denote the unit square, but we use the notation Ω when we think of it as a domain, while we use $\hat{K}$ when we think of it as a reference element), and for n a positive integer, let $\mathcal{U}_h$ be the uniform mesh of Ω into n^2 subsquares of side length $h = 1/n$. Given a subspace $\hat{\boldsymbol{V}}$ of $\boldsymbol{L}^2(\hat{K})$ we define

$$(2.2) \qquad \boldsymbol{V}_h = \{\boldsymbol{u} : \Omega \to \mathbb{R}^2 \mid \boldsymbol{u}|_K \in \boldsymbol{P}_{F_K}\hat{\boldsymbol{V}} \ \forall K \in \mathcal{U}_h\}.$$

In this definition, when we write $\boldsymbol{u}|_K \in \boldsymbol{P}_{F_K}\hat{\boldsymbol{V}}$ we mean only that $\boldsymbol{u}|_K$ agrees with a function in $\boldsymbol{P}_{F_K}\hat{\boldsymbol{V}}$ almost everywhere, and so do not impose any interelement continuity. Then we have the following approximation results.

THEOREM 2.1. *Let $\hat{\boldsymbol{V}}$ be a finite-dimensional subspace of $\boldsymbol{L}^2(\hat{K})$ and r be a nonnegative integer. The following conditions are equivalent:*

(i) *There is a constant C such that $\inf_{\boldsymbol{v}\in\boldsymbol{V}_h}\|\boldsymbol{u} - \boldsymbol{v}\|_{L^2(\Omega)} \le Ch^{r+1}|\boldsymbol{u}|_{H^{r+1}(\Omega)}$ for all $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(\Omega)$.*

(ii) *$\inf_{\boldsymbol{v}\in\boldsymbol{V}_h}\|\boldsymbol{u} - \boldsymbol{v}\|_{L^2(\Omega)} = o(h^r)$ for all $\boldsymbol{u} \in \boldsymbol{\mathcal{P}}_r(\Omega)$.*

(iii) *$\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(\hat{K})$.*

THEOREM 2.2. *Let $\hat{\boldsymbol{V}}$ be a finite-dimensional subspace of $\boldsymbol{L}^2(\hat{K})$ and r be a nonnegative integer. The following conditions are equivalent:*

(i) *There is a constant C such that*

$$\inf_{\boldsymbol{v}\in\boldsymbol{V}_h}\|\mathrm{div}\,\boldsymbol{u} - \mathrm{div}\,\boldsymbol{v}\|_{L^2(\Omega)} \le Ch^{r+1}|\mathrm{div}\,\boldsymbol{u}|_{H^{r+1}(\Omega)}$$

for all $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(\Omega)$ with $\mathrm{div}\,\boldsymbol{u} \in H^{r+1}(\Omega)$.

(ii) *$\inf_{\boldsymbol{v}\in\boldsymbol{V}_h}\|\mathrm{div}\,\boldsymbol{u} - \mathrm{div}\,\boldsymbol{v}\|_{L^2(\Omega)} = o(h^r)$ for all $\boldsymbol{u}$ with $\mathrm{div}\,\boldsymbol{u} \in \mathcal{P}_r(\Omega)$.*

(iii) *$\hat{\mathrm{div}}\,\hat{\boldsymbol{V}} \supseteq \mathcal{P}_r(\hat{K})$.*

Remark. Since we do not impose interelement continuity in the definition of $\boldsymbol{V}_h$, in Theorem 2.2 $\mathrm{div}\,\boldsymbol{v}$ should be interpreted as the divergence applied elementwise to $\boldsymbol{v} \in \boldsymbol{V}_h$.

3. A necessary condition for optimal approximation of vector fields on general quadrilateral meshes. In this section, we determine the properties of the finite element approximating spaces that are necessary for order $r + 1$ L^2-approximation of a vector field and its divergence on quadrilateral meshes. The construction of the finite element spaces proceeds as in the previous section. We start with the reference shape functions, a finite-dimensional space $\hat{\boldsymbol{V}}$ of vector fields on the unit square $\hat{K} = [0,1] \times [0,1]$ (typically $\hat{\boldsymbol{V}}$ consists of polynomials). Given an arbitrary convex quadrilateral K and a bilinear isomorphism $\boldsymbol{F}_K$ of the reference element $\hat{K}$ onto K, the shape functions on K are then taken to be $\boldsymbol{P}_{F_K}\hat{\boldsymbol{V}}$. (Note that there are eight possible choices for the bilinear isomorphism $\boldsymbol{F}_K$, but the space $\boldsymbol{P}_{F_K}\hat{\boldsymbol{V}}$ does not depend on the particular choice whenever $\hat{\boldsymbol{V}}$ is invariant under the symmetries of the square, which is usually the case in practice. When that is not the case, which we shall allow, it is necessary to specify not only the elements K but for each a choice of bilinear isomorphism from the reference element to K.) Finally, given a quadrilateral mesh $\mathcal{T}$ of a two-dimensional domain Ω, we can then construct the space of vector fields $\boldsymbol{V}(\mathcal{T})$ consisting of functions on Ω which belong to $\boldsymbol{P}_{F_K}\hat{\boldsymbol{V}}$ when restricted to a generic quadrilateral $K \in \mathcal{T}$.

It follows from the results of the previous section that if we consider the sequence $\mathcal{T}_h = \mathcal{U}_h$ of meshes of the unit square into congruent subsquares of side length $h = 1/n$,

then the approximation estimate

$$(3.1) \qquad \inf_{\boldsymbol{v}\in\boldsymbol{V}(\mathcal{T}_h)} \|\boldsymbol{u} - \boldsymbol{v}\|_{L^2(\Omega)} = o(h^r) \quad \forall\, \boldsymbol{u} \in \boldsymbol{\mathcal{P}}_r(\Omega)$$

is valid only if $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(\hat{K})$ and the estimate

$$(3.2) \qquad \inf_{\boldsymbol{v}\in\boldsymbol{V}(\mathcal{T}_h)} \|\operatorname{div}\boldsymbol{u} - \operatorname{div}\boldsymbol{v}\|_{L^2(\Omega)} = o(h^r) \quad \forall\, \boldsymbol{u} \text{ with } \operatorname{div}\boldsymbol{u} \in \mathcal{P}_r(\Omega)$$

is valid only if $\hat{\operatorname{div}}(\hat{\boldsymbol{V}}) \supseteq \mathcal{P}_r(\hat{K})$. In this section we show that for these estimates to hold for more general quadrilateral mesh sequences $\mathcal{T}_h$, stronger conditions on $\hat{\boldsymbol{V}}$ are required.

Before stating the main results of this section, we briefly recall a measure for the shape regularity of a convex quadrilateral K, cf. [7, A.2, pp. 104–105] or [13]. From the quadrilateral K we obtain four triangles by the four possible choices of three vertices from the vertices of K, and we define ρ_K as the smallest diameter of the inscribed circles to these four triangles. The *shape constant* of K is then $\sigma_K := h_K/\rho_K$, where $h_K = \operatorname{diam}(K)$. A bound on σ_K implies a bound on the ratio of any two sides of K and also a bound away from 0 and π for its angles (and conversely such bounds imply an upper bound on σ_K). It also implies bounds on the Lipschitz constant of $h_K^{-1}\boldsymbol{F}_K$ and its inverse. The shape constant of a mesh $\mathcal{T}_h$ consisting of convex quadrilaterals is then defined to be the supremum of the shape constants σ_K for $K \in \mathcal{T}_h$, and a family $\mathcal{T}_h$ of such meshes is called shape-regular if the shape constants for the meshes can be uniformly bounded.

The following two theorems give necessary conditions on the shape functions in order to ensure estimates like (3.1) and (3.2) on arbitrary quadrilateral mesh sequences. The spaces $\boldsymbol{\mathcal{S}}_r$ and $\mathcal{R}_r$ were defined in section 1.

THEOREM 3.1. *Suppose that the estimate* (3.1) *holds whenever* $\mathcal{T}_h$ *is a shape-regular sequence of quadrilateral meshes of a two-dimensional domain* Ω. *Then* $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{S}}_r$.

THEOREM 3.2. *Suppose that the estimate* (3.2) *holds whenever* $\mathcal{T}_h$ *is a shape-regular sequence of quadrilateral meshes of a two-dimensional domain* Ω. *Then* $\hat{\operatorname{div}}\,\hat{\boldsymbol{V}} \supseteq \mathcal{R}_r$.

In order to establish the theorems, we shall make use of two results analogous to Theorem 4 of [1]. To state these results, we introduce some specific bilinear mappings. For $\alpha > 0$, let $\boldsymbol{F}^\alpha$ and $\boldsymbol{G}^\alpha$ denote the mappings

$$(3.3) \qquad \boldsymbol{F}^\alpha(\hat{\boldsymbol{x}}) = (\hat{x}_1, (\alpha + \hat{x}_1)\hat{x}_2), \qquad \boldsymbol{G}^\alpha(\hat{\boldsymbol{x}}) = \boldsymbol{F}^\alpha(\hat{x}_2, \hat{x}_1),$$

each of which maps the unit square $\hat{K}$ to the quadrilateral K^α with vertices $(0,0)$, $(1,0)$, $(1, \alpha+1)$, and $(0, \alpha)$.

LEMMA 3.3. *Let* $\hat{\boldsymbol{V}}$ *be a space of vector fields on* $\hat{K}$ *such that* $\boldsymbol{P}_F\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(\boldsymbol{F}(\hat{K}))$ *when* $\boldsymbol{F}$ *is any of the four bilinear isomorphisms* $\boldsymbol{F}^1$, $\boldsymbol{F}^2$, $\boldsymbol{G}^1$, *and* $\boldsymbol{G}^2$. *Then* $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{S}}_r$.

LEMMA 3.4. *Let* $\hat{\boldsymbol{V}}$ *be a space of vector fields on* $\hat{K}$ *such that* $\operatorname{div}\boldsymbol{P}_F\hat{\boldsymbol{V}} \supseteq \mathcal{P}_r(\boldsymbol{F}(\hat{K}))$ *when* $\boldsymbol{F}$ *is any of the four bilinear isomorphisms* $\boldsymbol{F}^1$, $\boldsymbol{F}^2$, $\boldsymbol{G}^1$, *and* $\boldsymbol{G}^2$. *Then* $\hat{\operatorname{div}}\,\hat{\boldsymbol{V}} \supseteq \mathcal{R}_r$.

We postpone the proof of these lemmas to the end of the section. Now, based on Lemma 3.3 and Theorem 2.1, we establish Theorem 3.1.

Proof of Theorem 3.1. To establish the theorem, we assume that $\hat{\boldsymbol{V}} \not\supseteq \boldsymbol{\mathcal{S}}_r$ and exhibit a sequence $\mathcal{T}_h$ of shape regular meshes ($h = 1, 1/2, 1/3, \dots$) of the unit square
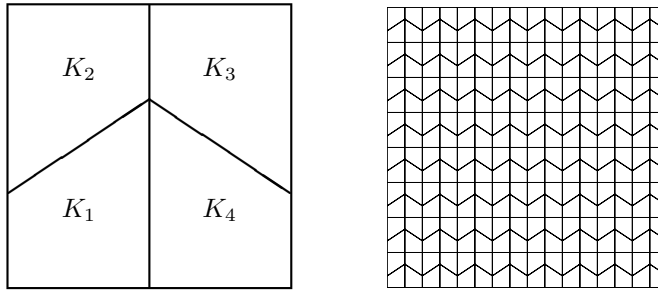
FIG. 1. (a) *The mesh* $\mathfrak{T}_1$ *of the unit square into four trapezoids.* (b) *The mesh* $\mathfrak{T}_h$ *(here* $h = 1/8$*) composed of translated dilates of* $\mathfrak{T}_1$.

for which the estimate (3.1) does not hold. We know, by Lemma 3.3, that for either $\alpha = 1$ or $\alpha = 2$ either $\boldsymbol{P}_{F^\alpha}\hat{\boldsymbol{V}}$ or $\boldsymbol{P}_{G^\alpha}\hat{\boldsymbol{V}}$ does not contain $\boldsymbol{\mathcal{P}}_r(K^\alpha)$. We fix this value of α and, without loss of generality, suppose that

$$(3.4) \qquad\qquad \boldsymbol{P}_{F^\alpha}\hat{\boldsymbol{V}} \not\supseteq \boldsymbol{\mathcal{P}}_r(K^\alpha).$$

Set $\beta = \alpha/(1 + 2\alpha)$. As shown in Figure 1(a), we define a mesh $\mathfrak{T}_1$ consisting of four congruent elements $K_1, \ldots, K_4$, with the vertices of K_1 given by $(0, 0)$, $(1/2, 0)$, $(1/2, 1-\beta)$, and $(0, \beta)$. For $h = 1/n$, we construct the mesh $\mathfrak{T}_h$ by partitioning the unit square into n^2 subsquares K and meshing each subsquare K with the mesh obtained by applying $\boldsymbol{F}_K$, given by (2.1), to $\mathfrak{T}_1$ as shown in Figure 1(b). For each element T of the mesh $\mathfrak{T}_h$ there is a natural way to construct a bilinear mapping $\boldsymbol{F}$ from the unit square onto T based on the mapping $\boldsymbol{F}^\alpha$. The first step is to compose $\boldsymbol{F}^\alpha$ with the linear isomorphism $\boldsymbol{E}(\boldsymbol{x}) = (x_1/2, x_2/(1 + 2\alpha))$ to obtain a bilinear map from the unit square onto the trapezoid K_1. Composing further with the natural isometries of K_1 onto K_2, K_3, and K_4, we obtain bilinear maps $\boldsymbol{F}_j$ from the unit square onto each of the trapezoids K_j, $j = 1, \ldots, 4$. Finally, further composition with the map $\boldsymbol{F}_K$ (consisting of dilation and translation) taking the unit square onto the subsquare K containing T, defines a bilinear diffeomorphism of the unit square onto T.

Having specified the mesh $\mathfrak{T}_h$ and a bilinear map from the unit square onto each element of the mesh, we have determined the space $\boldsymbol{V}(\mathfrak{T}_h)$ based on the shape functions in $\hat{\boldsymbol{V}}$. We need to show that the estimate (3.1) does not hold. To do so, we observe that $\boldsymbol{V}(\mathfrak{T}_h)$ coincides precisely with the space $\boldsymbol{V}_h$ constructed at the start of section 2 (see (2.2)) if we use $\boldsymbol{V}(\mathfrak{T}_1)$ as the space of shape functions on the unit square to begin the construction. This observation is easily verified in view of the composition property (1.2) of the Piola transform. Thus we may invoke Theorem 2.1 to conclude that (3.1) does not hold if we can show that $\boldsymbol{V}(\mathfrak{T}_1) \not\supseteq \boldsymbol{\mathcal{P}}_r(\hat{K})$. Now, by construction, the functions in $\boldsymbol{V}(\mathfrak{T}_1)$ restrict to functions in $\boldsymbol{P}_{F_1}\hat{\boldsymbol{V}}$ on $K_1 = \boldsymbol{F}_1\hat{K}$, so it is enough to show that $\boldsymbol{P}_{F_1}\hat{\boldsymbol{V}} \not\supseteq \boldsymbol{\mathcal{P}}_r(K_1)$. But $\boldsymbol{F}_1 = \boldsymbol{E} \circ \boldsymbol{F}^\alpha$ and, hence, $\boldsymbol{P}_{F_1}\hat{\boldsymbol{V}} = \boldsymbol{P}_E(\boldsymbol{P}_{F^\alpha}\hat{\boldsymbol{V}})$. Now $\boldsymbol{E}$ is a linear isomorphism of K^α onto K_1, and so $\boldsymbol{P}_E$ is a linear isomorphism of $\boldsymbol{\mathcal{P}}_r(K^\alpha)$ onto $\boldsymbol{\mathcal{P}}_r(K_1)$. Thus $\boldsymbol{P}_{F_1}\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(K_1)$ if and only if $\boldsymbol{P}_{F^\alpha}\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(K^\alpha)$ and so the theorem is complete in view of (3.4). $\qquad\square$

Proof of Theorem 3.2. The proof is essentially identical to the preceding one, except that Lemma 3.4 and Theorem 2.2 are used in place of Lemma 3.3 and Theorem 2.1. $\qquad\square$

Before turning to the proof of Lemmas 3.3 and 3.4, we draw some implications from Theorems 3.1 and 3.2 for the approximation properties of the extensions of stan-

dard finite element subspaces of $\boldsymbol{H}(\mathrm{div}, \Omega)$ from rectangular meshes to quadrilateral meshes. By definition, $\boldsymbol{\mathcal{S}}_r \subseteq \boldsymbol{\mathcal{RT}}_r$, so Theorem 3.1 does not contradict the possibility that the Raviart–Thomas space of index r achieves order $r + 1$ approximation in $\boldsymbol{L}^2$ on quadrilateral meshes, just as for rectangular meshes. This is indeed the case (see the discussion in section 1). But $\mathrm{div}\,\boldsymbol{\mathcal{RT}}_r = \mathcal{Q}_r$ which contains $\mathcal{R}_{r-1}$ but not $\mathcal{R}_r$. Thus we may conclude from Theorem 3.2 that the best possible order of approximation to the divergence in L^2 for the Raviart–Thomas space of index r is only r on quadrilateral meshes, one degree lower than for rectangular meshes, and, in particular, there is no convergence for $r = 0$. (This lower order is achieved, as discussed in section 1.) In contrast to the Raviart–Thomas spaces, for the Brezzi–Douglas–Marini and Brezzi–Douglas–Fortin–Marini spaces there is a loss of L^2-approximation order on quadrilateral meshes. Both $\boldsymbol{\mathcal{BDM}}_r$ and $\boldsymbol{\mathcal{BDFM}}_{r+1}$ contain $\boldsymbol{\mathcal{P}}_r$, which is enough to ensure order $r + 1$ approximation in $\boldsymbol{L}^2$ on rectangular meshes. However, it is easy to check that $\boldsymbol{\mathcal{BDM}}_r$ contains $\boldsymbol{\mathcal{S}}_{\lfloor (r-1)/2 \rfloor}$ but not $\boldsymbol{\mathcal{S}}_{\lfloor (r+1)/2 \rfloor}$ so that the best possible order of approximation for the Brezzi–Douglas–Marini space of index r on general quadrilateral meshes is $\lfloor (r+1)/2 \rfloor$, a substantial loss of accuracy in comparison to the rectangular case. For the divergence, we have $\hat{\mathrm{div}}\,\boldsymbol{\mathcal{BDM}}_r = \mathcal{P}_{r-1}(\hat{K})$ which contains $\mathcal{R}_{\lfloor (r-2)/2 \rfloor}$ but not $\mathcal{R}_{\lfloor r/2 \rfloor}$. Therefore the best possible order of approximation for the divergence for the Brezzi–Douglas–Marini space of index r on general quadrilateral meshes is $\mathcal{R}_{\lfloor r/2 \rfloor}$. Similarly, the best possible order of L^2-approximation for the Brezzi–Douglas–Fortin–Marini space of index $r + 1$ on general quadrilateral meshes is $\lfloor (r + 2)/2 \rfloor$, while since $\hat{\mathrm{div}}\,\boldsymbol{\mathcal{BDFM}}_{r+1} = \mathcal{P}_r(\hat{K})$, the best possible rate for the divergence is $\lfloor (r + 1)/2 \rfloor$. We specifically note that in the lowest index cases, namely when $\hat{\boldsymbol{V}} = \boldsymbol{\mathcal{RT}}_0$, $\boldsymbol{\mathcal{BDM}}_1$, or $\boldsymbol{\mathcal{BDFM}}_1$ (which is identical to $\boldsymbol{\mathcal{RT}}_0$), *the best approximation in $\boldsymbol{H}(\mathrm{div}, \Omega)$ does not converge in $\boldsymbol{H}(\mathrm{div}, \Omega)$ for general quadrilateral mesh sequences.* Section 8 of this paper contains a numerical confirmation of this result.

We conclude this section with the proofs of Lemmas 3.3 and 3.4.

Proof of Lemma 3.3. By hypothesis $\boldsymbol{P}_F \hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(\boldsymbol{F}(\hat{K}))$ or, equivalently, $\hat{\boldsymbol{V}} \supseteq \boldsymbol{P}_F^{-1}[\boldsymbol{\mathcal{P}}_r(\boldsymbol{F}(\hat{K}))]$, for $\boldsymbol{F} = \boldsymbol{F}^1$, $\boldsymbol{F}^2$, $\boldsymbol{G}^1$, and $\boldsymbol{G}^2$. Thus it is sufficient to prove that

$$(3.5) \quad \boldsymbol{\mathcal{S}}_r \subseteq \Sigma_r := \boldsymbol{P}_{F^1}^{-1}[\boldsymbol{\mathcal{P}}_r(K^1)] + \boldsymbol{P}_{F^2}^{-1}[\boldsymbol{\mathcal{P}}_r(K^2)] + \boldsymbol{P}_{G^1}^{-1}[\boldsymbol{\mathcal{P}}_r(K^1)] + \boldsymbol{P}_{G^2}^{-1}[\boldsymbol{\mathcal{P}}_r(K^2)].$$

We will prove this using induction on r.

Now for any diffeomorphism $\boldsymbol{F} : \hat{K} \to K$ and any $\boldsymbol{u} : K \to \mathbb{R}^2$, we have, directly from the definition of the Piola transform, that

$$(3.6) \qquad (\boldsymbol{P}_F^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) = J\boldsymbol{F}(\hat{\boldsymbol{x}})D\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1}\boldsymbol{u}(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial F_2}{\partial \hat{x}_2}(\hat{\boldsymbol{x}}) & -\dfrac{\partial F_1}{\partial \hat{x}_2}(\hat{\boldsymbol{x}}) \\[2mm] -\dfrac{\partial F_2}{\partial \hat{x}_1}(\hat{\boldsymbol{x}}) & \dfrac{\partial F_1}{\partial \hat{x}_1}(\hat{\boldsymbol{x}}) \end{pmatrix} \boldsymbol{u}(\boldsymbol{x}).$$

Specializing to the case where $\boldsymbol{F} = \boldsymbol{F}^\alpha$ or $\boldsymbol{G}^\alpha$ given by (3.3), we have

$$(\boldsymbol{P}_{F^\alpha}^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) = \begin{pmatrix} \alpha + \hat{x}_1 & 0 \\ -\hat{x}_2 & 1 \end{pmatrix}\boldsymbol{u}(\boldsymbol{x}), \qquad (\boldsymbol{P}_{G^\alpha}^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) = \begin{pmatrix} \hat{x}_1 & -1 \\ -\alpha - \hat{x}_2 & 0 \end{pmatrix}\boldsymbol{u}(\boldsymbol{x}).$$

Thus, when $u(\boldsymbol{x})$ is the constant vector field $(1, 0)$, $(\boldsymbol{P}_{F^1}^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) = (1 + \hat{x}_1, -\hat{x}_2)$, and when $u(\boldsymbol{x}) \equiv (0, 1)$, $(\boldsymbol{P}_{F^1}^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) = (0, 1)$ and $(\boldsymbol{P}_{G^1}^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) = (-1, 0)$. These three vector fields span $\boldsymbol{\mathcal{S}}_0$, which establishes (3.5) in the case $r = 0$.

Suppose now that $\boldsymbol{\mathcal{S}}_{r-1} \subseteq \Sigma_{r-1}$ for some $r \geq 1$. To complete the induction we need to show that $\boldsymbol{\mathcal{S}}_r \subseteq \Sigma_r$. Now $\boldsymbol{\mathcal{S}}_r$ is spanned by $\boldsymbol{\mathcal{S}}_{r-1}$ plus the $4r + 4$ additional

vector fields

$$
\begin{array}{llll}
(\hat{x}_1^i \hat{x}_2^r, 0) & \text{and} & (0, \hat{x}_1^r \hat{x}_2^i), & 0 \le i \le r, \\
(\hat{x}_1^{r+1} \hat{x}_2^j, 0) & \text{and} & (0, \hat{x}_1^j \hat{x}_2^{r+1}), & 0 \le j \le r-1, \\
(\hat{x}_1^r \hat{x}_2^{r-1}, 0) & \text{and} & (\hat{x}_1^{r+1} \hat{x}_2^r, -\hat{x}_1^r \hat{x}_2^{r+1}).
\end{array}
$$

Pick $0 \le i \le r$, and set $\boldsymbol{F} = \boldsymbol{G}^\alpha$ and $\boldsymbol{u}(\boldsymbol{x}) = (0, -x_1^{r-i} x_2^i) \in \boldsymbol{\mathcal{P}}_r(K^\alpha)$. Note that $\boldsymbol{x} = \boldsymbol{G}^\alpha \hat{\boldsymbol{x}} = (\hat{x}_2, (\alpha + \hat{x}_2)\hat{x}_1)$. Then

$$
(\boldsymbol{P}_{G^\alpha}^{-1} \boldsymbol{u})(\hat{\boldsymbol{x}}) = (x_1^{r-i} x_2^i, 0) = (\hat{x}_2^{r-i} (\alpha + \hat{x}_2)^i \hat{x}_1^i, 0) = (\hat{x}_1^i \hat{x}_2^r, 0) + i\alpha(\hat{x}_1^i \hat{x}_2^{r-1}, 0)
$$
$$
\pmod{\boldsymbol{\mathcal{S}}_{r-1}}.
$$

Since $\boldsymbol{\mathcal{S}}_{r-1} \subseteq \boldsymbol{\Sigma}_r$ by the inductive hypothesis, and since we may take both $\alpha = 1$ and $\alpha = 2$, we conclude that $(\hat{x}_1^i \hat{x}_2^r, 0) \in \boldsymbol{\Sigma}_r$ (for $0 \le i \le r$) and also that $(\hat{x}_1^i \hat{x}_2^{r-1}, 0) \in \boldsymbol{\Sigma}_r$.

In a similar way, setting $\boldsymbol{F} = \boldsymbol{F}^\alpha$ and $\boldsymbol{u}(\boldsymbol{x}) = (0, x_1^{r-i} x_2^i)$, we conclude that $(0, \hat{x}_1^r \hat{x}_2^i) \in \boldsymbol{\Sigma}_r$, $0 \le i \le r$. The choice $\boldsymbol{F} = \boldsymbol{F}^\alpha$ and $\boldsymbol{u}(\boldsymbol{x}) = (x_1^{r-j} x_2^j, 0)$ together with the fact that $\boldsymbol{\Sigma}_r \supseteq \mathcal{Q}_r \times \mathcal{Q}_r$, which is a consequence of the proof thus far, implies that $(\hat{x}_1^{r+1} \hat{x}_2^j, 0) \in \boldsymbol{\Sigma}_r$ for $0 \le j \le r-1$. The choice $\boldsymbol{F} = \boldsymbol{G}^\alpha$ with the same choice of $\boldsymbol{u}$ similarly implies that $(0, \hat{x}_1^j \hat{x}_2^{r+1}) \in \boldsymbol{\Sigma}_r$ for $0 \le j \le r-1$.

Finally, with $\boldsymbol{u}(\boldsymbol{x}) = (x_2^r, 0)$, we find that $(\boldsymbol{P}_{F^1}^{-1} \boldsymbol{u})(\hat{\boldsymbol{x}}) = (\hat{x}_1^{r+1} \hat{x}_2^r, -\hat{x}_1^r \hat{x}_2^{r+1})$ $\pmod{\mathcal{Q}_r \times \mathcal{Q}_r}$, which completes the proof of (3.5) and so the lemma. $\square$

Proof of Lemma 3.4. The hypothesis is that $\text{div}\, \boldsymbol{P}_F \hat{V} \supseteq \mathcal{P}_r(\boldsymbol{F}(\hat{K}))$ for $\boldsymbol{F} = \boldsymbol{F}^1$, $\boldsymbol{F}^2$, $\boldsymbol{G}^1$, and $\boldsymbol{G}^2$. Now $\hat{\text{div}}\, \hat{\boldsymbol{u}}(\hat{\boldsymbol{x}}) = J\boldsymbol{F}(\hat{\boldsymbol{x}}) \text{div}(\boldsymbol{P}_F \hat{\boldsymbol{u}})(\boldsymbol{F}\hat{\boldsymbol{x}})$, so $\hat{\text{div}}\, \hat{V}$ contains all functions on $\hat{K}$ of the form $\hat{\boldsymbol{x}} \mapsto J\boldsymbol{F}(\hat{\boldsymbol{x}}) p(\boldsymbol{F}\hat{\boldsymbol{x}})$ with $p \in \mathcal{P}_r(\boldsymbol{F}(\hat{K}))$ and $\boldsymbol{F} \in \{\boldsymbol{F}^1, \boldsymbol{F}^2, \boldsymbol{G}^1, \boldsymbol{G}^2\}$. To prove the lemma, it suffices to show that the span of such functions, call it Σ_r, contains $\mathcal{R}_r$. Note that $J\boldsymbol{F}^\alpha(\hat{\boldsymbol{x}}) = \alpha + \hat{x}_1$ and $J\boldsymbol{G}^\alpha(\hat{\boldsymbol{x}}) = -\alpha - \hat{x}_2$.

For $r = 0$, we take $p \equiv 1$ and $\boldsymbol{F} = \boldsymbol{F}^1$, $\boldsymbol{F}^2$, and $\boldsymbol{G}^1$, and find that Σ_r contains $1 + \hat{x}_1$, $2 + \hat{x}_1$, and $-1 - \hat{x}_2$. These three functions span $\mathcal{R}_0$, so $\Sigma_0 \supseteq \mathcal{R}_0$.

We continue the proof that $\Sigma_r \supseteq \mathcal{R}_r$ by induction on r. Now $\mathcal{R}_r$ is the span of $\mathcal{R}_{r-1}$ and the $2r + 3$ additional functions $\hat{x}_1^{r+1} \hat{x}_2^i$ and $\hat{x}_1^i \hat{x}_2^{r+1}$, $0 \le i \le r$, and $\hat{x}_1^r \hat{x}_2^r$. Taking $p(\boldsymbol{x}) = x_1^{r-i} x_2^i$ and $\boldsymbol{F} = \boldsymbol{F}^\alpha$ we find that the function $\hat{\boldsymbol{x}} \mapsto \hat{x}_1^{r-i} (\alpha + \hat{x}_1)^{i+1} \hat{x}_2^i$ belongs to Σ_r. Modulo $\mathcal{R}_{r-1}$ (which is contained in Σ_r by the inductive hypothesis), this is equal to the function $\hat{\boldsymbol{x}} \mapsto \hat{x}_1^{r+1} \hat{x}_2^i + (i+1)\alpha \hat{x}_1^r \hat{x}_2^i$. Using both $\alpha = 1$ and 2, we conclude that $\hat{x}_1^{r+1} \hat{x}_2^i$ belongs to Σ_r for $0 \le i \le r$ and that $\hat{x}_1^r \hat{x}_2^r$ does as well. The same choice of p with $\boldsymbol{F} = \boldsymbol{G}^\alpha$ shows that Σ_r contains the functions $\hat{x}_1^i \hat{x}_2^{r+1}$, $0 \le i \le r$, and completes the proof. $\square$

4. Sufficient conditions for optimal order approximation. In this section we show that the necessary conditions we have obtained in the previous section are also sufficient for approximation of order $r + 1$ in $\boldsymbol{L}^2$ and $\boldsymbol{H}(\text{div}, \Omega)$. To state this more precisely, we recall the construction of projection operators for $\boldsymbol{H}(\text{div})$ finite elements. We suppose that we are given a bounded projection $\hat{\boldsymbol{\pi}} : \boldsymbol{H}^{r+1}(\hat{K}) \to \hat{V}$ (typically this operator is specified via a unisolvent set of degrees of freedom for $\hat{V}$). We then define the corresponding projection $\boldsymbol{\pi}_K : \boldsymbol{H}^{r+1}(K) \to \boldsymbol{P}_F \hat{V}$ for an arbitrary element $K = \boldsymbol{F}(\hat{K})$ via the Piola transform, as expressed in this commuting diagram:

$$
\begin{array}{ccc}
\boldsymbol{H}^{r+1}(\hat{K}) & \xrightarrow{\hat{\boldsymbol{\pi}}} & \hat{V} \\
{\scriptstyle P_F} \downarrow & & \downarrow {\scriptstyle P_F} \\
\boldsymbol{H}^{r+1}(K) & \xrightarrow[\boldsymbol{\pi}_K]{} & \boldsymbol{P}_F \hat{V}
\end{array}.
$$

That is, $\boldsymbol{\pi}_K = \boldsymbol{P}_F \circ \hat{\boldsymbol{\pi}} \circ \boldsymbol{P}_F^{-1}$. Finally a global projection operator $\boldsymbol{\pi}_h : \boldsymbol{H}^{r+1}(\Omega) \to \boldsymbol{V}(\mathfrak{T}_h)$ is defined piecewise: $(\boldsymbol{\pi}_h \boldsymbol{u})|_K = \boldsymbol{\pi}_K(\boldsymbol{u}|_K)$. (The degrees of freedom used to define $\hat{\boldsymbol{\pi}}$ will determine the degree of interelement continuity enjoyed by $\boldsymbol{\pi}_h \boldsymbol{u}$. In particular, for the standard $\boldsymbol{H}(\mathrm{div})$ finite element spaces discussed previously, the degrees of freedom ensure that on any edge $\hat{e}$ of $\hat{K}$, $(\hat{\boldsymbol{\pi}}\boldsymbol{u}) \cdot \hat{\boldsymbol{n}}$ on $\hat{e}$ depends only on $\boldsymbol{u} \cdot \boldsymbol{n}$ on $\hat{e}$. From this it results that $\boldsymbol{\pi}_h \boldsymbol{u} \in \boldsymbol{H}(\mathrm{div})$.)

The following two theorems contain the main results of this section.

THEOREM 4.1. *Let $\hat{\boldsymbol{\pi}} : \boldsymbol{H}^{r+1}(\hat{K}) \to \hat{\boldsymbol{V}}$ be a bounded projection operator. Given a quadrilateral mesh $\mathfrak{T}_h$ of a domain Ω, let $\boldsymbol{\pi}_h : \boldsymbol{H}^{r+1}(\Omega) \to \boldsymbol{V}(\mathfrak{T}_h)$ be defined as above. Suppose that $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{S}}_r$. Then there exists a constant C depending only on the bound for $\hat{\boldsymbol{\pi}}$ and on the shape regularity of $\mathfrak{T}_h$, such that*

$$(4.1) \qquad \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^{r+1} |\boldsymbol{u}|_{H^{r+1}(\Omega)}$$

for all $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(\Omega)$.

THEOREM 4.2. *Let $\hat{\boldsymbol{\pi}} : \boldsymbol{H}^{r+1}(\hat{K}) \to \hat{\boldsymbol{V}}$ be a bounded projection operator. Given a quadrilateral mesh $\mathfrak{T}_h$ of a domain Ω, let $\boldsymbol{\pi}_h : \boldsymbol{H}^{r+1}(\Omega) \to \boldsymbol{V}(\mathfrak{T}_h)$ be defined as above. Suppose that $\hat{\mathrm{div}}\, \hat{\boldsymbol{V}} \supseteq \mathcal{R}_r$. Suppose also that there exists a bounded projection operator $\hat{\Pi} : H^{r+1}(\hat{K}) \to \hat{\mathrm{div}}\, \hat{\boldsymbol{V}}$ such that*

$$(4.2) \qquad \hat{\mathrm{div}}\, \hat{\boldsymbol{\pi}}\hat{\boldsymbol{u}} = \hat{\Pi}\, \hat{\mathrm{div}}\, \hat{\boldsymbol{u}} \quad \forall \hat{\boldsymbol{u}} \in \boldsymbol{H}^{r+1}(\hat{K})\ \text{with}\ \hat{\mathrm{div}}\, \hat{\boldsymbol{u}} \in H^{r+1}(\hat{K}).$$

Then there exists a constant C depending only on the bounds for $\hat{\boldsymbol{\pi}}$ and $\hat{\Pi}$ and on the shape regularity of $\mathfrak{T}_h$, such that

$$(4.3) \qquad \|\mathrm{div}\, \boldsymbol{u} - \mathrm{div}\, \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^{r+1} |\mathrm{div}\, \boldsymbol{u}|_{H^{r+1}(\Omega)}$$

for all $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(\Omega)$ with $\mathrm{div}\, \boldsymbol{u} \in H^{r+1}(\Omega)$.

Remarks. 1. It follows immediately that if the hypotheses of both theorems are met, then $\boldsymbol{\pi}_h$ furnishes order $r + 1$ approximation in $\boldsymbol{H}(\mathrm{div}, \Omega)$:

$$\|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{\boldsymbol{H}(\mathrm{div}, \Omega)} \leq C h^{r+1}(|\boldsymbol{u}|_{H^{r+1}(\Omega)} + |\mathrm{div}\, \boldsymbol{u}|_{H^{r+1}(\Omega)})$$

for all $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(\Omega)$ with $\mathrm{div}\, \boldsymbol{u} \in H^{r+1}(\Omega)$.

2. The commutativity hypothesis involving the projection $\hat{\Pi}$ plays a major role in the theory of $\boldsymbol{H}(\mathrm{div}, \Omega)$ finite elements. It is satisfied in the case of the Raviart–Thomas, Brezzi–Douglas–Marini, and Brezzi–Douglas–Fortin–Marini elements, as well as for the new elements introduced in the next sections, with $\hat{\Pi}$ equal to the L^2 projection onto $\hat{\mathrm{div}}\, \hat{\boldsymbol{V}}$.

3. When applied to the Raviart–Thomas elements of index r, Theorem 4.1 gives

$$\|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^{r+1} |\boldsymbol{u}|_{H^{r+1}(\Omega)}$$

and Theorem 4.2 gives

$$\|\mathrm{div}\, \boldsymbol{u} - \mathrm{div}\, \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^r |\mathrm{div}\, \boldsymbol{u}|_{H^{r+1}(\Omega)}.$$

The latter estimate is proved in [12], but the former estimate appears to be new. It improves on the estimate given in [12]:

$$\|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{L^2(\Omega)} \leq C h^{r+1}[|\boldsymbol{u}|_{H^{r+1}(\Omega)} + h |\mathrm{div}\, \boldsymbol{u}|_{H^{r+1}(\Omega)}].$$

The proofs of the theorems depend on the following two lemmas which are strengthened converses of Lemmas 3.3 and 3.4.

LEMMA 4.3. *Let $\hat{\boldsymbol{V}}$ be a space of vector fields on $\hat{K}$ containing $\boldsymbol{\mathcal{S}}_r$. Then $\boldsymbol{P}_F\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(K)$ for all bilinear isomorphisms $\boldsymbol{F}$ of $\hat{K}$ onto convex quadrilaterals $K = \boldsymbol{F}(\hat{K})$.*

Proof. It is sufficient to show that $\boldsymbol{\mathcal{S}}_r \supseteq \boldsymbol{P}_F^{-1}[\boldsymbol{\mathcal{P}}_r(K)]$, since then the hypothesis $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{S}}_r$ implies that

$$\boldsymbol{P}_F\hat{\boldsymbol{V}} \supseteq \boldsymbol{P}_F\boldsymbol{\mathcal{S}}_r \supseteq \boldsymbol{P}_F\boldsymbol{P}_F^{-1}[\boldsymbol{\mathcal{P}}_r(K)] = \boldsymbol{\mathcal{P}}_r(K).$$

Now (3.6) tells us that

$$\boldsymbol{P}_F^{-1}\boldsymbol{u} = \begin{pmatrix} \partial F_2/\partial\hat{x}_2 & -\partial F_1/\partial\hat{x}_2 \\ -\partial F_2/\partial\hat{x}_1 & \partial F_1/\partial\hat{x}_1 \end{pmatrix} (\boldsymbol{u} \circ \boldsymbol{F}).$$

Since $\boldsymbol{u} \in \boldsymbol{\mathcal{P}}_r(K)$ and $\boldsymbol{F}$ is bilinear, $\boldsymbol{u} \circ \boldsymbol{F} \in \boldsymbol{\mathcal{Q}}_r(\hat{K})$. Also, again in view of the bilinearity of $\boldsymbol{F}$, the matrix appearing in this equation is the sum of a constant matrix field and one of the form $(\hat{x}_1, -\hat{x}_2)^T(a_2, -a_1)$ (where $a_i \in \mathbb{R}$ is the coefficient of $\hat{x}_1\hat{x}_2$ in F_i). It follows immediately that $\boldsymbol{P}_F^{-1}\boldsymbol{u} \in \boldsymbol{\mathcal{S}}_r$. ☐

LEMMA 4.4. *Let $\hat{\boldsymbol{V}}$ be a space of vector fields on $\hat{K}$ such that $\hat{\mathrm{div}}\,\hat{\boldsymbol{V}} \supseteq \mathcal{R}_r$. Then $\mathrm{div}\,\boldsymbol{P}_F\hat{\boldsymbol{V}} \supseteq \mathcal{P}_r(K)$ for all bilinear isomorphisms $\boldsymbol{F}$ of $\hat{K}$ onto convex quadrilaterals $K = \boldsymbol{F}(\hat{K})$.*

Proof. Let $p \in \mathcal{P}_r(K)$ be arbitrary. Choose any $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ such that $\mathrm{div}\,\boldsymbol{u} = p$. From the identity

$$(\hat{\mathrm{div}}\,\boldsymbol{P}_F^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) = J\boldsymbol{F}(\hat{\boldsymbol{x}})(\mathrm{div}\,\boldsymbol{u})(\boldsymbol{x}),$$

we have $\hat{\mathrm{div}}\,\boldsymbol{P}_F^{-1}\boldsymbol{u} = J\boldsymbol{F} \cdot (p \circ \boldsymbol{F})$. Now $p \in \mathcal{P}_r(K)$ and $\boldsymbol{F}$ is bilinear, so $p \circ \boldsymbol{F}$ belongs to $\mathcal{Q}_r(\hat{K})$ and $J\boldsymbol{F}$ is linear. Thus $\hat{q} := \hat{\mathrm{div}}\,\boldsymbol{P}_F^{-1}\boldsymbol{u} \in \mathcal{R}_r$.

Invoking the hypothesis that $\mathcal{R}_r \subseteq \hat{\mathrm{div}}\,\hat{\boldsymbol{V}}$, we can find $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$ such that $\hat{\mathrm{div}}\,\hat{\boldsymbol{v}} = \hat{q}$. Then

$$\begin{aligned} p(\boldsymbol{x}) = \mathrm{div}\,\boldsymbol{u}(\boldsymbol{x}) &= J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1}(\hat{\mathrm{div}}\,\boldsymbol{P}_F^{-1}\boldsymbol{u})(\hat{\boldsymbol{x}}) \\ &= J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1}\hat{q}(\hat{\boldsymbol{x}}) = J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1}\,\hat{\mathrm{div}}\,\hat{\boldsymbol{v}}(\hat{\boldsymbol{x}}) = \mathrm{div}\,\boldsymbol{P}_F\hat{\boldsymbol{v}}(\boldsymbol{x}). \end{aligned}$$

This shows that $p \in \mathrm{div}\,\boldsymbol{P}_F\hat{\boldsymbol{V}}$ as required. ☐

Proof of Theorem 4.1. We will show that if $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{S}}_r$ and K is any convex quadrilateral, then

(4.4) $$\|\boldsymbol{u} - \boldsymbol{\pi}_K\boldsymbol{u}\|_{L^2(K)} \leq Ch_K^{r+1}|\boldsymbol{u}|_{H^{r+1}(K)} \quad \forall\boldsymbol{u} \in \boldsymbol{H}^{r+1}(K),$$

where $h_K = \mathrm{diam}(K)$ and the constant C depends only on $\hat{\boldsymbol{\pi}}$ and the shape constant for K. The theorem follows easily by squaring both sides and summing over the elements.

We establish (4.4) in two steps. First we prove it under the additional assumption that $h_K = 1$, and then we use a simple scaling argument to obtain it for arbitrary K.

For the first part we use the Bramble–Hilbert lemma. In view of Lemma 4.3 and the fact that $\hat{\boldsymbol{\pi}}$ is a projection onto $\hat{\boldsymbol{V}}$, it follows that $\boldsymbol{\pi}_K\boldsymbol{u} = \boldsymbol{u}$ for all $\boldsymbol{u} \in \boldsymbol{\mathcal{P}}_r(K)$. Now under the assumption that $h_K = 1$, the Piola transform $\boldsymbol{P}_{F_K}$ is bounded and invertible both from $\boldsymbol{L}^2(\hat{K})$ to $\boldsymbol{L}^2(K)$ and from $\boldsymbol{H}^{r+1}(\hat{K})$ to $\boldsymbol{H}^{r+1}(K)$ with bounds in both norms depending only on the shape constant. A similar statement holds for $\boldsymbol{P}_{F_K}^{-1}$.

Since $\hat{\boldsymbol{\pi}}$ is bounded from $\boldsymbol{H}^{r+1}(\hat{K})$ to $\boldsymbol{L}^2(\hat{K})$, it follows that $\boldsymbol{\pi}_K = \boldsymbol{P}_{F_K} \circ \hat{\boldsymbol{\pi}} \circ \boldsymbol{P}_{F_K}^{-1}$ is bounded from $\boldsymbol{H}^{r+1}(K)$ to $\boldsymbol{L}^2(K)$ with bound depending only on the bound for $\hat{\boldsymbol{\pi}}$ and the shape constant for K. The map $\boldsymbol{u} \mapsto \boldsymbol{u} - \boldsymbol{\pi}_K \boldsymbol{u}$ is then similarly bounded and moreover vanishes on $\boldsymbol{\mathcal{P}}_r(K)$. Therefore,

$$\|\boldsymbol{u} - \boldsymbol{\pi}_K \boldsymbol{u}\|_{L^2(K)} \leq \|\boldsymbol{I} - \boldsymbol{\pi}_K\|_{\mathcal{L}(\boldsymbol{H}^{r+1}(K), \boldsymbol{L}^2(K))} \inf_{\boldsymbol{p} \in \boldsymbol{\mathcal{P}}_r(K)} \|\boldsymbol{u} - \boldsymbol{p}\|_{H^{r+1}(K)}.$$

Now the Bramble–Hilbert lemma states that the last infimum can be bounded by $c|\boldsymbol{u}|_{H^{r+1}(K)}$, where c depends only on r and the shape regularity of K (see, e.g., [2, Lemma 4.3.8]). The estimate (4.4) then follows for $h_K = 1$ with $C = c\|\boldsymbol{I} - \boldsymbol{\pi}_K\|_{\mathcal{L}(\boldsymbol{H}^{r+1}(K), \ll^2(K))}$.

To complete the proof, let K be an arbitrary convex quadrilateral, and denote by $\boldsymbol{M} : K \to \tilde{K} := h_K^{-1} K$ the dilation $\boldsymbol{M}(\boldsymbol{x}) = h_K^{-1} \boldsymbol{x}$. Then the bilinear maps $\boldsymbol{F}_K$ and $\boldsymbol{F}_{\tilde{K}}$ of the reference element $\hat{K}$ onto K and $\tilde{K}$, respectively, are related by the equation $\boldsymbol{F}_{\tilde{K}} = \boldsymbol{M} \circ \boldsymbol{F}_K$, from which it follows easily that $\boldsymbol{\pi}_{\tilde{K}} = \boldsymbol{P}_M \circ \boldsymbol{\pi}_K \circ \boldsymbol{P}_M^{-1}$. Of course, $\boldsymbol{P}_M$ has a very simple form:

$$\boldsymbol{P}_M \boldsymbol{u}(\tilde{\boldsymbol{x}}) = h_K \boldsymbol{u}(h_K \tilde{\boldsymbol{x}}).$$

Now for any $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(K)$, let $\tilde{\boldsymbol{u}} = \boldsymbol{P}_M \boldsymbol{u} \in \boldsymbol{H}^{r+1}(\tilde{K})$. It is then easy to check that

$$\|\boldsymbol{u} - \boldsymbol{\pi}_K \boldsymbol{u}\|_{L^2(K)} = \|\boldsymbol{P}_M^{-1}(\tilde{\boldsymbol{u}} - \boldsymbol{\pi}_{\tilde{K}} \tilde{\boldsymbol{u}})\|_{L^2(K)} = \|\tilde{\boldsymbol{u}} - \boldsymbol{\pi}_{\tilde{K}} \tilde{\boldsymbol{u}}\|_{L^2(\tilde{K})}$$
$$\leq C|\tilde{\boldsymbol{u}}|_{H^{r+1}(\tilde{K})} = Ch_K^{r+1}|\boldsymbol{u}|_{H^{r+1}(K)},$$

where we obtained the inequality from the already established result for elements of unit diameter. ⬜

Proof of Theorem 4.2. As for the previous theorem, it suffices to prove a local result:

$$(4.5) \qquad \| \operatorname{div} \boldsymbol{u} - \operatorname{div} \boldsymbol{\pi}_K \boldsymbol{u} \|_{L^2(K)} \leq Ch_K^{r+1} |\operatorname{div} \boldsymbol{u}|_{H^{r+1}(K)}$$
$$\forall \boldsymbol{u} \in \boldsymbol{H}^{r+1}(K) \text{ with } \operatorname{div} \boldsymbol{u} \in H^{r+1}(K),$$

where C depends only on the bounds for $\hat{\boldsymbol{\pi}}$ and $\hat{\Pi}$ and the shape constant of K.

Define $\Lambda_K : L^2(K) \to L^2(K)$ by

$$(4.6) \qquad \Lambda_K p(\boldsymbol{x}) = J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1} \hat{\Pi}[J\boldsymbol{F} \cdot (p \circ \boldsymbol{F})](\hat{\boldsymbol{x}}),$$

i.e., $\Lambda_K p = \{J\boldsymbol{F}^{-1} \cdot \hat{\Pi}[J\boldsymbol{F} \cdot (p \circ \boldsymbol{F})]\} \circ \boldsymbol{F}^{-1}$. Then

$$\operatorname{div} \boldsymbol{\pi}_K \boldsymbol{u}(\boldsymbol{x}) = \operatorname{div}(\boldsymbol{P}_{F_K} \hat{\boldsymbol{\pi}} \boldsymbol{P}_{F_K}^{-1} \boldsymbol{u})(\boldsymbol{x}) = J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1} \hat{\operatorname{div}}(\hat{\boldsymbol{\pi}} \boldsymbol{P}_{F_K}^{-1} \boldsymbol{u})(\hat{\boldsymbol{x}})$$
$$= J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1} \hat{\Pi}(\hat{\operatorname{div}} \boldsymbol{P}_{F_K}^{-1} \boldsymbol{u})(\hat{\boldsymbol{x}}) = J\boldsymbol{F}(\hat{\boldsymbol{x}})^{-1} \hat{\Pi}[J\boldsymbol{F} \cdot (\operatorname{div} \boldsymbol{u}) \circ \boldsymbol{F}](\hat{\boldsymbol{x}}).$$

That is, $\operatorname{div} \boldsymbol{\pi}_K \boldsymbol{u} = \Lambda_K(\operatorname{div} \boldsymbol{u})$. Thus,

$$\| \operatorname{div} \boldsymbol{u} - \operatorname{div} \boldsymbol{\pi}_K \boldsymbol{u} \|_{L^2(K)} = \| \operatorname{div} \boldsymbol{u} - \Lambda_K(\operatorname{div} \boldsymbol{u}) \|_{L^2(K)}$$

and (4.5) will hold if we can prove that

$$(4.7) \qquad \|p - \Lambda_K p\|_{L^2(K)} \leq Ch_K^{r+1} |p|_{H^{r+1}(K)} \quad \forall p \in H^{r+1}(K).$$

The proof of (4.7) is again given first in the case of elements of unit diameter. Then Λ_K is bounded uniformly from $H^{r+1}(K)$ to $L^2(K)$ for elements K with uniformly
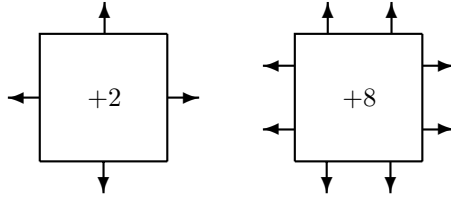
Fig. 2. *Element diagrams indicating the degrees of freedom for* $\boldsymbol{ABF}_0$ *and* $\boldsymbol{ABF}_1$.

bounded shape constant. Now, as noted in the proof of Lemma 4.4, if $p \in \mathcal{P}_r(K)$, then $J\boldsymbol{F} \cdot (p \circ \boldsymbol{F}) \in \mathcal{R}_r \subseteq \hat{\mathrm{div}}\,\hat{\boldsymbol{V}}$. Since $\hat{\Pi}$ is a projection onto $\hat{\mathrm{div}}\,\hat{\boldsymbol{V}}$, it follows that $\Lambda_K p = p$ for $p \in \mathcal{P}_r(K)$. Thus the Bramble–Hilbert lemma implies (4.7) under the restriction $h_K = 1$. To extend to elements of arbitrary diameter, we again use a dilation.   □

5. Construction of spaces with optimal order $\boldsymbol{H}(\mathrm{div}, \Omega)$ approxima-tion. We have previously shown that none of the standard finite element approximations of $\boldsymbol{H}(\mathrm{div}, \Omega)$ (i.e., the Raviart–Thomas, Brezzi–Douglas–Marini, or Brezzi–Douglas–Fortin–Marini spaces) maintain the same order of approximation on general convex quadrilaterals as they do on rectangles. In this section, we use the conditions determined in the previous sections to construct finite element subspaces of $\boldsymbol{H}(\mathrm{div}, \Omega)$ which do have this property. To obtain approximation of order $r + 1$ in $\boldsymbol{H}(\mathrm{div}, \Omega)$ on general convex quadrilaterals, we require that the space of reference shape functions $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{S}}_r$ and $\hat{\mathrm{div}}\,\hat{\boldsymbol{V}} \supseteq \mathcal{R}_r$. A space with this property is $\boldsymbol{ABF}_r := \mathcal{P}_{r+2,r}(\hat{K}) \times \mathcal{P}_{r,r+2}(\hat{K})$, for which $\hat{\mathrm{div}}\,\boldsymbol{ABF}_r = \mathcal{R}_r$.

As degrees of freedom for $\boldsymbol{ABF}_r$ on the reference element, we take

$$(5.1) \qquad \int_{\hat{e}} \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}}\hat{q}\,d\hat{s}, \quad \hat{q} \in \mathcal{P}_r(\hat{e}) \text{ for each edge } \hat{e} \text{ of } \hat{K}$$

$$(5.2) \qquad \int_{\hat{K}} \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{\phi}}\,d\hat{\boldsymbol{x}}, \quad \hat{\boldsymbol{\phi}} \in \mathcal{P}_{r-1,r}(\hat{K}) \times \mathcal{P}_{r,r-1}(\hat{K}),$$

$$(5.3) \qquad \int_{\hat{K}} \hat{\mathrm{div}}\,\hat{\boldsymbol{u}}\hat{x}_1^{r+1}\hat{x}_2^i\,d\hat{\boldsymbol{x}}, \quad \int_{\hat{K}} \hat{\mathrm{div}}\,\hat{\boldsymbol{u}}\hat{x}_1^i\hat{x}_2^{r+1}\,d\hat{\boldsymbol{x}}, \quad i = 0, \ldots, r.$$

Note that (5.1) and (5.2) are the standard degrees of freedom for the Raviart–Thomas elements on the reference square. In all we have specified $4(r+1)+2r(r+1)+2(r+1) = 2(r+3)(r+1) = \dim \boldsymbol{ABF}_r$ degrees of freedom. Since the new degrees of freedom, with respect to the standard Raviart–Thomas elements, are local, we remark that the implementation of the new space $\boldsymbol{ABF}_r$ should not be more expensive than that of $\boldsymbol{RT}_r$. Figure 2 indicates the degrees of freedom for the first two cases $r = 0$ and 1.

In order to see that these choices of $\hat{\boldsymbol{V}}$ and degrees of freedom determine a finite element subspace of $\boldsymbol{H}(\mathrm{div}, \Omega)$, we need to show that the degrees of freedom are unisolvent, and that if the degrees of freedom on an edge $\hat{e}$ vanish, then $\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}}$ vanishes on e (this will ensure that the assembled finite element space belongs to $\boldsymbol{H}(\mathrm{div}, \Omega)$). The second point is immediate. On any edge $\hat{e}$ of $\hat{K}$, $\boldsymbol{u} \cdot \boldsymbol{n} \in \mathcal{P}_r(\hat{e})$, so the vanishing of the degrees of freedom (5.1) associated to $\hat{e}$ does indeed ensure that $\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}} \equiv 0$.

We now verify unisolvence by showing that if $\hat{\boldsymbol{u}} \in \boldsymbol{ABF}_r$ and all the quantities (5.1)–(5.3) vanish, then $\hat{\boldsymbol{u}} = 0$. If $\hat{q} \in \mathcal{Q}_r(\hat{K})$, then $\hat{q}|_{\hat{e}} \in \mathcal{P}_r(\hat{e})$ for any edge $\hat{e}$ of

$\hat{K}$, and $\hat{\nabla}\hat{q} \in \mathcal{P}_{r-1,r}(\hat{K}) \times \mathcal{P}_{r,r-1}(\hat{K})$. Therefore,

$$\int_{\hat{K}} \hat{\mathrm{div}}\, \hat{\boldsymbol{u}}\hat{q}\, d\hat{\boldsymbol{x}} = \int_{\partial\hat{K}} \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}}\hat{q}\, ds - \int_{\hat{K}} \hat{\boldsymbol{u}} \cdot \hat{\nabla}\hat{q}\, d\hat{\boldsymbol{x}} = 0, \quad \hat{q} \in \mathcal{Q}_r(\hat{K}).$$

In view of (5.3) we then have that

$$\int_{\hat{K}} \hat{\mathrm{div}}\, \hat{\boldsymbol{u}}\hat{q}\, d\hat{\boldsymbol{x}} = 0, \quad \hat{q} \in \mathcal{R}_r.$$

Since $\hat{\mathrm{div}}\, \hat{\boldsymbol{u}} \in \mathcal{R}_r$ we conclude that $\hat{\mathrm{div}}\, \hat{\boldsymbol{u}} = 0$. Now we may write

$$\hat{\boldsymbol{u}} = \sum_{i=0}^{r} [a_i(\hat{x}_1^{r+2}\hat{x}_2^i, 0) + b_i(0, \hat{x}_1^i\hat{x}_2^{r+2})] + \hat{\boldsymbol{v}}$$

with $\hat{\boldsymbol{v}} \in \mathcal{RT}_r$. Since

$$0 = \hat{\mathrm{div}}\, \hat{\boldsymbol{u}} = \sum_{i=1}^{r} (r+2)(a_i\hat{x}_1^{r+1}\hat{x}_2^i + b_i\hat{x}_1^i\hat{x}_2^{r+1}) + \hat{\mathrm{div}}\, \hat{\boldsymbol{v}},$$

and $\hat{\mathrm{div}}\, \hat{\boldsymbol{v}} \in \mathcal{Q}_r$, it follows that $a_i = b_i = 0$ and so $\hat{\boldsymbol{u}} = \hat{\boldsymbol{v}} \in \mathcal{RT}_r$. Since (5.1), (5.2) are unisolvent degrees of freedom for $\mathcal{RT}_r$ [4, Proposition III.3.4], we conclude that $\hat{\boldsymbol{u}} = 0$.

We also note that a small variant of the first part of this argument establishes the commutativity property (4.2) with $\hat{\boldsymbol{\pi}} : \boldsymbol{H}^1(\hat{K}) \to \mathcal{ABF}_r$ the projection determined by the degrees of freedom (5.1)–(5.3) and $\hat{\Pi}$ the L^2-projection onto $\mathcal{R}_r = \hat{\mathrm{div}}\, \mathcal{ABF}_r$. Thus, all the hypotheses of Theorems 4.1 and 4.2 are satisfied and the estimates (4.1) and (4.3) hold on general quadrilateral meshes for finite element spaces based on $\mathcal{ABF}_r$.

6. Application to mixed finite element methods. One of the main applications of finite element subspaces of $\boldsymbol{H}(\mathrm{div}, \Omega)$ is to the approximation of second order elliptic boundary value problems by mixed finite element methods. For the model problem $\Delta p = f$ in Ω, $p = 0$ on $\partial\Omega$, the mixed formulation is the following: Find $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ and $p \in L^2(\Omega)$ such that

$$(\boldsymbol{u}, \boldsymbol{v}) + (p, \mathrm{div}\, \boldsymbol{v}) = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega),$$
$$(\mathrm{div}\, \boldsymbol{u}, q) = (f, q) \quad \forall q \in L^2(\Omega),$$

where $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product. For $\boldsymbol{S}_h \subseteq \boldsymbol{H}(\mathrm{div}, \Omega)$ and $W_h \subseteq L^2(\Omega)$, the mixed finite element approximation seeks $\boldsymbol{u}_h \in \boldsymbol{S}_h$ and $p_h \in W_h$ such that

$$(\boldsymbol{u}_h, \boldsymbol{v}) + (p_h, \mathrm{div}\, \boldsymbol{v}) = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{S}_h,$$
$$(\mathrm{div}\, \boldsymbol{u}_h, q) = (f, q) \quad \forall q \in W_h.$$

The pair $(\boldsymbol{S}_h, W_h)$ is said to be stable if the following conditions are satisfied:

$$(6.1) \quad (\boldsymbol{v}, \boldsymbol{v}) \geq c\|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div}, \Omega)}^2 \quad \forall \boldsymbol{v} \in \boldsymbol{Z}_h = \{\boldsymbol{v} \in \boldsymbol{S}_h : (\mathrm{div}\, \boldsymbol{v}, q) = 0\ \forall q \in W_h\},$$

$$(6.2) \quad \sup_{\boldsymbol{v} \in \boldsymbol{S}_h} \frac{(\mathrm{div}\, \boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div}, \Omega)}} \geq c\|q\|_{L^2(\Omega)} \quad \forall q \in W_h.$$

By Brezzi's theorem [3], if $(\boldsymbol{S}_h, W_h)$ is a stable pair, then the quasioptimality estimate

(6.3)
$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}(\mathrm{div},\Omega)} + \|p - p_h\|_{L^2(\Omega)} \leq C\Big(\inf_{\boldsymbol{v} \in \boldsymbol{S}_h} \|\boldsymbol{u} - \boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div},\Omega)} + \inf_{q \in W_h} \|p - q\|_{L^2(\Omega)} \Big)$$

holds with C depending only on Ω and the constant c entering into the stability conditions.

For the space $\boldsymbol{S}_h$ we will take $\boldsymbol{V}(\mathcal{T}_h) \cap \boldsymbol{H}(\mathrm{div}, \Omega)$, where $\mathcal{T}_h$ is an arbitrary quadrilateral mesh and $\boldsymbol{V}(\mathcal{T}_h)$ is constructed as described at the start of section 3 starting from a space of reference shape functions $\hat{\boldsymbol{V}}$ on the unit square. To specify the corresponding space W_h, we first define a space of reference shape functions $\hat{W} = \hat{\mathrm{div}}\,\hat{\boldsymbol{V}}$, next define the space of shape functions on K by $W_K = \{\hat{w} \circ \boldsymbol{F}_K^{-1} \mid \hat{w} \in \hat{W}\}$, and then set

$$W_h = \{w \in L^2(\Omega) \mid w|_K \in W_K\}.$$

Now suppose that $\hat{\boldsymbol{V}}$ is any one of the previously considered spaces $\boldsymbol{\mathcal{RT}}_r$, $\boldsymbol{\mathcal{BDM}}_r$, $\boldsymbol{\mathcal{BDFM}}_{r+1}$, or $\boldsymbol{\mathcal{ABF}}_r$. Associated with each of these spaces is a unisolvent set of degrees of freedom. These are given in (5.1) and (5.2) for $\boldsymbol{\mathcal{RT}}_r$, by (5.1)–(5.3) for $\boldsymbol{\mathcal{ABF}}_r$, and, for $\boldsymbol{\mathcal{BDM}}_r$ and $\boldsymbol{\mathcal{BDFM}}_{r+1}$, by (5.1) and $\int_{\hat{K}} \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{\phi}}\, d\hat{\boldsymbol{x}}$ with $\hat{\boldsymbol{\phi}}$ in $\boldsymbol{\mathcal{P}}_{r-2}(\hat{K})$ or $\boldsymbol{\mathcal{P}}_{r-1}(\hat{K})$, respectively. These degrees of freedom determine the projection $\hat{\boldsymbol{\pi}} : \boldsymbol{H}^1(\hat{K}) \to \hat{\boldsymbol{V}}$ and then, by the construction described at the start of section 4, the projection $\boldsymbol{\pi}_h : \boldsymbol{H}^1(\Omega) \to \boldsymbol{S}_h$. Moreover, the degrees of freedom ensure the commutativity property (4.2) where $\hat{\Pi}$ is the $L^2(\hat{K})$ projection onto $\hat{W}$. From these observations it is straightforward to derive the stability conditions (6.1) and (6.2), as we shall now do.

Given $\boldsymbol{v} \in \boldsymbol{Z}_h$ and $K \in \mathcal{T}_h$, let $\hat{\boldsymbol{v}} = \boldsymbol{P}_{F_K}^{-1}(\boldsymbol{v}|_K) \in \hat{\boldsymbol{V}}$, $\hat{q} = \hat{\mathrm{div}}\,\hat{\boldsymbol{v}} \in \hat{W}$, and $q = \hat{q} \circ \boldsymbol{F}_K^{-1} \in W_K$. Then $(\mathrm{div}\,\boldsymbol{v}, q)_{L^2(K)} = 0$ (because we can extend q to Ω by zero and obtain a function in W_h and $\mathrm{div}\,\boldsymbol{v}$ is orthogonal to W_h since $\boldsymbol{v} \in \boldsymbol{Z}_h$). But $(\mathrm{div}\,\boldsymbol{v}, q)_{L^2(K)} = (\hat{\mathrm{div}}\,\hat{\boldsymbol{v}}, \hat{q})_{L^2(\hat{K})} = \|\hat{\mathrm{div}}\,\hat{\boldsymbol{v}}\|_{L^2(K)}^2$, so $\hat{\mathrm{div}}\,\hat{\boldsymbol{v}} = 0$ and therefore $\mathrm{div}\,\boldsymbol{v} = [(J\boldsymbol{F}_K)^{-1}\,\hat{\mathrm{div}}\,\hat{\boldsymbol{v}}] \circ \boldsymbol{F}_K^{-1} = 0$. Thus, if $\boldsymbol{v} \in \boldsymbol{Z}_h$, then $\mathrm{div}\,\boldsymbol{v} = 0$, and (6.1) follows immediately with $c = 1$.

To prove (6.2), we shall show that for any given $q \in W_h$ there exists $\boldsymbol{v} \in \boldsymbol{S}_h$ with

(6.4)
$$(\mathrm{div}\,\boldsymbol{v}, q) = \|q\|_{L^2(\Omega)}$$

and

(6.5)
$$\|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div},\Omega)} \leq C\|q\|_{L^2(\Omega)}.$$

As usual, we start by noting that there exists $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ with $\mathrm{div}\,\boldsymbol{u} = q$ and $\|\boldsymbol{u}\|_{H^1(\Omega)} \leq C\|q\|_{L^2(\Omega)}$ and letting $\boldsymbol{v} = \boldsymbol{\pi}_h\boldsymbol{u}$. Now $(\mathrm{div}\,\boldsymbol{\pi}_h\boldsymbol{u}, q) = (\mathrm{div}\,\boldsymbol{u}, q)$ whenever $q \in W_h$, as follows directly from the construction of $\boldsymbol{\pi}_h$, the commutativity property (4.2), and the properties of the Piola transform. Therefore (6.4) holds. To prove (6.5) we note that in each case $\hat{\boldsymbol{V}} \supseteq \boldsymbol{S}_0$, so Theorem 4.1 gives the estimate $\|\boldsymbol{u} - \boldsymbol{\pi}_h\boldsymbol{u}\|_{L^2(\Omega)} \leq Ch\|\boldsymbol{u}\|_{H^1(\Omega)}$, and so, by the triangle inequality, $\|\boldsymbol{v}\|_{L^2(\Omega)} \leq C\|q\|_{L^2(\Omega)}$. Also, on any element K, $\mathrm{div}\,\boldsymbol{v} = \mathrm{div}\,\boldsymbol{\pi}_K\boldsymbol{u} = \Lambda_K(\mathrm{div}\,\boldsymbol{u}) = \Lambda_K q$, where Λ_K is defined by (4.6), which implies that $\|\mathrm{div}\,\boldsymbol{v}\|_{L^2(\Omega)} \leq C\|q\|_{L^2(\Omega)}$. This establishes (6.5) and completes the proof of stability.

Remark. Note that we do not have $W_h = \operatorname{div} \boldsymbol{S}_h$ on general quadrilateral meshes, although this is the case on rectangular meshes. With that choice of W_h it would be easy to prove (6.1) but the proof of (6.2) would not be clear.

We now turn our attention to error estimates for mixed methods. Having established stability, we can combine the quasioptimality estimate (6.3) with the bounds for the approximation error given by Theorems 4.1 and 4.2 (and Theorem 1 of [1] for the approximation error for p) to obtain error bounds. For the $\boldsymbol{\mathcal{ABF}}_r$ method this gives

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}(\operatorname{div},\Omega)} + \|p - p_h\|_{L^2(\Omega)} \leq Ch^{r+1}(|u|_{H^{r+1}(\Omega)} + |\operatorname{div}\boldsymbol{u}|_{H^{r+1}(\Omega)} + |p|_{H^{r+1}(\Omega)}).$$

But for the $\boldsymbol{\mathcal{RT}}_r$ method it gives only an $O(h^r)$ bound, and no convergence at all for $r = 0$, because of the decreased approximation for the divergence (and the approximation orders are even lower for $\boldsymbol{\mathcal{BDM}}_r$ and $\boldsymbol{\mathcal{BDFM}}_{r+1}$).

It is possible to improve on this by following the approach of [6] and [5], as we now do. First we define $\Pi_K : L^2(K) \to W_K$ by $\Pi_K p = (\hat{\Pi}\hat{p}) \circ \boldsymbol{F}_K^{-1}$ with $\hat{p} = p \circ \boldsymbol{F}_K$, and then we define $\Pi_h : L^2(\Omega) \to W_h$ by $\Pi_h p|_K = \Pi_K(p|_K)$. It follows that $(p - \Pi_K p, \operatorname{div} \boldsymbol{v})_{L^2(K)} = (\hat{p} - \hat{\Pi}\hat{p}, \hat{\operatorname{div}} \boldsymbol{P}_{F_K}^{-1}\boldsymbol{v})_{L^2(K)}$, so

$$(p - \Pi_h p, \operatorname{div} \boldsymbol{v}) = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{S}_h.$$

We then have the following error estimates.

THEOREM 6.1.

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)} \leq \|\boldsymbol{u} - \boldsymbol{\pi}_h\boldsymbol{u}\|_{L^2(\Omega)},$$
$$\|\operatorname{div}\boldsymbol{u}_h\|_{L^2(\Omega)} \leq C\|\operatorname{div}\boldsymbol{u}\|_{L^2(\Omega)},$$
$$\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)} \leq C\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{\pi}_h\boldsymbol{u})\|_{L^2(\Omega)},$$
$$\|\Pi_h p - p_h\|_{L^2(\Omega)}^2 = (\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{U} - \boldsymbol{\pi}_h\boldsymbol{U}) + (\operatorname{div}[\boldsymbol{u} - \boldsymbol{u}_h], P - \Pi_h P),$$

where P is the solution to the Dirichlet problem $-\Delta P = \Pi_h p - p_h$ in Ω, $P = 0$ on $\partial\Omega$ and $\boldsymbol{U} = \operatorname{grad} P$.

Proof. Using the error equations

$$(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}) + (p - p_h, \operatorname{div}\boldsymbol{v}) = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{S}_h, \qquad (\operatorname{div}[\boldsymbol{u} - \boldsymbol{u}_h], q) = 0 \quad \forall q \in W_h,$$

we obtain

$$(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\pi}_h\boldsymbol{u} - \boldsymbol{u}_h) = (p - p_h, \operatorname{div}[\boldsymbol{u}_h - \boldsymbol{\pi}_h\boldsymbol{u}]) = (\Pi_h p - p_h, \operatorname{div}[\boldsymbol{u}_h - \boldsymbol{\pi}_h\boldsymbol{u}])$$
$$= (\Pi_h p - p_h, \operatorname{div}[\boldsymbol{u}_h - \boldsymbol{u}]) = 0.$$

Hence, $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}^2 = (\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{\pi}_h\boldsymbol{u})$ and it easily follows that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)} \leq \|\boldsymbol{u} - \boldsymbol{\pi}_h\boldsymbol{u}\|_{L^2(\Omega)}.$$

To estimate $\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$, we observe that if $v \in \boldsymbol{S}_h$ and we define

$$q(\boldsymbol{x}) = \begin{cases} |J\boldsymbol{F}_K(\hat{\boldsymbol{x}})| \operatorname{div}\boldsymbol{v}(\boldsymbol{x}), & \boldsymbol{x} \in K, \\ 0, & \boldsymbol{x} \in \Omega \setminus K, \end{cases}$$

then $q \in W_h$. Therefore, from the error equation we have

$$(\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h), |J\boldsymbol{F}_K| \operatorname{div}\boldsymbol{v})_K = 0.$$

Choosing $\boldsymbol{v} = \boldsymbol{u}_h$, it easily follows that

$$\||J\boldsymbol{F}_K|^{1/2}\operatorname{div}\boldsymbol{u}_h\|_{L^2(K)} \leq \||J\boldsymbol{F}_K|^{1/2}\operatorname{div}\boldsymbol{u}\|_{L^2(K)},$$

and so $\|\operatorname{div}\boldsymbol{u}_h\|_{L^2(K)} \leq C\|\operatorname{div}\boldsymbol{u}\|_{L^2(K)}$ with C depending on the shape constant for K. Choosing $\boldsymbol{v} = \boldsymbol{\pi}_h\boldsymbol{u} - \boldsymbol{u}_h$, it also follows that

$$\||J\boldsymbol{F}_K|^{1/2}\operatorname{div}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(K)} \leq \||J\boldsymbol{F}_K|^{1/2}\operatorname{div}(\boldsymbol{u}-\boldsymbol{\pi}_h\boldsymbol{u})\|_{L^2(K)},$$

so $\|\operatorname{div}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(K)} \leq C\|\operatorname{div}(\boldsymbol{u}-\boldsymbol{\pi}_h\boldsymbol{u})\|_{L^2(K)}$. Summing over all quadrilaterals, we obtain

$$\|\operatorname{div}\boldsymbol{u}_h\|_{L^2(\Omega)} \leq C\|\operatorname{div}\boldsymbol{u}\|_{L^2(\Omega)}, \qquad \|\operatorname{div}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega)} \leq C\|\operatorname{div}(\boldsymbol{u}-\boldsymbol{\pi}_h\boldsymbol{u})\|_{L^2(\Omega)}.$$

To estimate $\|p - p_h\|_{L^2(\Omega)}$, we define P as the solution to the Dirichlet problem $\Delta P = \Pi_h p - p_h$ in Ω, $P = 0$ on $\partial\Omega$, and set $\boldsymbol{U} = \operatorname{grad}P$. Then

$$\begin{aligned}
\|\Pi_h p - p_h\|_{L^2(\Omega)}^2 &= (\operatorname{div}\boldsymbol{U}, \Pi_h p - p_h) = (\operatorname{div}\boldsymbol{\pi}_h\boldsymbol{U}, \Pi_h p - p_h) = -(\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{\pi}_h\boldsymbol{U}) \\
&= (\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{U}-\boldsymbol{\pi}_h\boldsymbol{U}) - (\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{U}) \\
&= (\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{U}-\boldsymbol{\pi}_h\boldsymbol{U}) + (\operatorname{div}[\boldsymbol{u}-\boldsymbol{u}_h], P) \\
&= (\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{U}-\boldsymbol{\pi}_h\boldsymbol{U}) + (\operatorname{div}[\boldsymbol{u}-\boldsymbol{u}_h], P - \Pi_h P). \qquad \square
\end{aligned}$$

To obtain order of convergence estimates, one needs to apply the approximation properties of a particular space. For the Raviart–Thomas elements of index r we obtain the following estimates.

THEOREM 6.2. *Suppose* $(\boldsymbol{u}_h, p_h)$ *is the mixed method approximation to* $(\boldsymbol{u}, p)$ *obtained when* $\hat{\boldsymbol{V}}$ *is the Raviart–Thomas reference space of index* r *and suppose that the domain* Ω *is convex. Then for* $p \in H^{r+2}(\Omega)$,

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)} \leq Ch^{r+1}\|\boldsymbol{u}\|_{H^{r+1}(\Omega)},$$

$$\|\operatorname{div}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega)} \leq Ch^{r}\|\operatorname{div}\boldsymbol{u}\|_{H^{r}(\Omega)},$$

$$\|p - p_h\|_{L^2(\Omega)} \leq \begin{cases} Ch^{r+1}\|p\|_{H^{r+1}(\Omega)} & (r \geq 1), \\ Ch\|p\|_{H^2(\Omega)} & (r = 0). \end{cases}$$

Proof. It follows from [7, section I.A.2] that $\|p - \Pi_h p\|_{L^2(\Omega)} \leq Ch^{k+1}\|p\|_{k+1,\Omega}$, $0 \leq k \leq r$, and it follows from Theorems 4.1 and 4.2 that, for $0 \leq k \leq r$,

$$\|\boldsymbol{u}-\boldsymbol{\pi}_h\boldsymbol{u}\|_{L^2(\Omega)} \leq Ch^{k+1}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}, \qquad \|\operatorname{div}[\boldsymbol{u}-\boldsymbol{\pi}_h\boldsymbol{u}]\|_{L^2(\Omega)} \leq Ch^{k}\|\operatorname{div}\boldsymbol{u}\|_{H^{k}(\Omega)}.$$

Inserting these results in Theorem 6.1, we immediately obtain the first two estimates of Theorem 6.2. From the last estimate of Theorem 6.1, we also obtain

$$\|\Pi_h p - p_h\|_{L^2(\Omega)} \leq C(h\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)} + h^{\min(1+r,2)}\|\operatorname{div}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega)}).$$

Here we have used elliptic regularity, which holds under the assumption that Ω is convex, to bound $\|\boldsymbol{U}\|_{H^1(\Omega)} = \|P\|_{H^2(\Omega)}$ by $\|\Pi_h p - p_h\|_{L^2(\Omega)}$. Hence, for $r \geq 1$ and $0 \leq k \leq r$, we obtain

$$\|\Pi_h p - p_h\|_{L^2(\Omega)} \leq Ch^{k+2}(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|\operatorname{div}\boldsymbol{u}\|_{H^{k}(\Omega)}) \leq Ch^{k+2}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}.$$

Choosing $k = r - 1$ and $k = r$, we obtain for $r \geq 1$

$$\|\Pi_h p - p_h\|_{L^2(\Omega)} \leq Ch^{r+1}\|\boldsymbol{u}\|_{H^{r}(\Omega)}, \qquad \|\Pi_h p - p_h\|_{L^2(\Omega)} \leq Ch^{r+2}\|\boldsymbol{u}\|_{H^{r+1}(\Omega)},$$

and for $r = 0$, $\|\Pi_h p - p_h\|_{L^2(\Omega)} \leq Ch\|\boldsymbol{u}\|_{1,\Omega}$. The final estimates of the theorem now follow directly by the triangle inequality. $\square$

7. Application to least squares methods. A standard finite element least squares approximation of the Dirichlet problem $\Delta p = f$ in Ω, $p = 0$ on $\partial\Omega$ seeks $p_h \in W_h \subseteq H_0^1(\Omega)$ and $\boldsymbol{u}_h \in \boldsymbol{S}_h \subseteq \boldsymbol{H}(\mathrm{div}, \Omega)$ minimizing

$$J(q, \boldsymbol{v}) = \|\boldsymbol{v} - \mathrm{grad}\, q\|_{L^2(\Omega)}^2 + \|\mathrm{div}\, \boldsymbol{v} + f\|_{L^2(\Omega)}^2$$

over $W_h \times \boldsymbol{S}_h$. For any choices of subspaces this satisfies the quasioptimality estimate (cf. [10])

$$\|p - p_h\|_{H^1(\Omega)} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}(\mathrm{div},\Omega)} \le C\Big(\inf_{p \in W_h} \|p - q\|_{H^1(\Omega)} + \inf_{\boldsymbol{v} \in \boldsymbol{S}_h} \|\boldsymbol{u} - \boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div},\Omega)}\Big).$$

If we take W_h to be the standard H^1 finite element space based on reference shape functions $\mathcal{Q}_{r+1}$ and use the $\boldsymbol{\mathcal{ABF}}_r$ space for $\boldsymbol{S}_h$, we immediately obtain

$$\|p - p_h\|_{H^1(\Omega)} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}(\mathrm{div},\Omega)} \le Ch^{r+1}(\|p\|_{H^{r+1}(\Omega)} + \|\boldsymbol{u}\|_{H^{r+1}(\Omega)} + \|\mathrm{div}\, \boldsymbol{u}\|_{H^{r+1}(\Omega)}).$$

However, the quasioptimality estimate suggests that if we choose the same W_h but use the $\boldsymbol{\mathcal{RT}}_r$ elements for $\boldsymbol{S}_h$, the lower rate of approximation of $\mathrm{div}\,\boldsymbol{u}$ may negatively influence the approximation of both variables.

Next, we use a duality argument to obtain a second estimate, which provides improved convergence for p in L^2 when the $\boldsymbol{\mathcal{ABF}}$ spaces are used, but again suggests difficulties for the $\boldsymbol{\mathcal{RT}}$ spaces. We shall henceforth assume that the domain Ω is convex so that we have 2-regularity for the Dirichlet problem for the Laplacian. Define $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ and $r \in H_0^1(\Omega)$ as solution of the dual problem

$$(7.1) \qquad \int_\Omega (\boldsymbol{w} - \nabla r) \cdot \boldsymbol{v}\, d\boldsymbol{x} + \int_\Omega \mathrm{div}\,\boldsymbol{w}\, \mathrm{div}\,\boldsymbol{v}\, d\boldsymbol{x} = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega),$$

$$(7.2) \qquad \int_\Omega (\boldsymbol{w} - \nabla r) \cdot \nabla q\, d\boldsymbol{x} = -\int_\Omega (p - p_h) q\, d\boldsymbol{x} \quad \forall q \in H_0^1(\Omega).$$

This problem has a unique solution, since if $p - p_h$ were to vanish, then we could take $\boldsymbol{v} = \boldsymbol{w}$ and $q = r$, subtract the equations, and conclude that $\boldsymbol{w} = \nabla r$, $\mathrm{div}\,\boldsymbol{w} = 0$ with $r \in H_0^1(\Omega)$, which implies that $\boldsymbol{w}$ and r vanish. For general $p - p_h$, the solution of the dual problem may be written as $\boldsymbol{w} = \nabla(r + g)$ where $g \in H^2(\Omega) \cap H_0^1(\Omega)$ satisfies $\Delta g = p - p_h$ and $r \in H^2(\Omega) \cap H_0^1(\Omega)$ satisfies $\Delta r = g - p + p_h$ (so $\mathrm{div}\,\boldsymbol{w} = g$). Note that $\|r\|_{H^2(\Omega)} + \|\boldsymbol{w}\|_{H^1(\Omega)} + \|\mathrm{div}\,\boldsymbol{w}\|_{H^2(\Omega)} \le C\|p - p_h\|_{L^2(\Omega)}$. Choosing $q = p - p_h$, $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}_h$, subtracting (7.2) from (7.1), and using the error equations

$$\int_\Omega (\boldsymbol{u} - \boldsymbol{u}_h - \nabla[p - p_h]) \cdot \boldsymbol{v}\, d\boldsymbol{x} + \int_\Omega \mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\, \mathrm{div}\,\boldsymbol{v}\, d\boldsymbol{x} = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{S}_h,$$

$$\int_\Omega (\boldsymbol{u} - \boldsymbol{u}_h - \nabla[p - p_h]) \cdot \nabla q\, d\boldsymbol{x} = 0 \quad \forall q \in W_h,$$

one obtains the estimate

$$\|p - p_h\|_{L^2(\Omega)}^2 \le C(\|r - r_I\|_{H^1(\Omega)} + \|\boldsymbol{w} - \boldsymbol{w}_I\|_{\boldsymbol{H}(\mathrm{div},\Omega)})(\|p - p_h\|_{H^1(\Omega)} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}(\mathrm{div},\Omega)})$$

for all $\boldsymbol{w}_I \in \boldsymbol{S}_h$ and $r_I \in W_h$. This estimate will furnish an improved order of convergence for p in L^2 as compared to H^1 if $\boldsymbol{S}_h$ has good approximation properties in $\boldsymbol{H}(\mathrm{div}, \Omega)$. For the $\boldsymbol{\mathcal{ABF}}_r$ space (still with W_h based on $\mathcal{Q}_{r+1}$) we obtain

$$\|p - p_h\|_{L^2(\Omega)} \le Ch^{r+2}(\|p\|_{H^{r+1}(\Omega)} + \|\boldsymbol{u}\|_{H^{r+r}(\Omega)} + \|\mathrm{div}\,\boldsymbol{u}\|_{H^{r+1}(\Omega)}).$$

But for the $\boldsymbol{\mathcal{RT}}_0$ space we obtain no convergence whatsoever. Numerical computations reported in the next section verify these findings for both the scalar and the vector variable: with W_h taken to be the usual four node H^1 elements based on $\mathcal{Q}_1$ and $\boldsymbol{S}_h$ based on $\boldsymbol{\mathcal{ABF}}_0$, we obtain convergence of order 1 for $\boldsymbol{u}$ in $\boldsymbol{H}(\mathrm{div}, \Omega)$ and of order 2 for p in $L^2(\Omega)$, but if we use $\boldsymbol{\mathcal{RT}}_0$ elements instead there is no L^2 convergence for $\boldsymbol{u}$ or p.

The numerical computations of the next section also exhibit second order convergence for $\mathrm{div}\,\boldsymbol{u}$ in $\boldsymbol{L}^2(\Omega)$ when approximated by the $\boldsymbol{\mathcal{ABF}}_0$ method on square meshes. We close this section by showing that

$$\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)} = O(h^{r+2})$$

when the $\boldsymbol{\mathcal{ABF}}_r$ elements are used on rectangular meshes. Now define $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ and $r \in H_0^1(\Omega)$ by

$$\int_\Omega (\boldsymbol{w} - \nabla r) \cdot \boldsymbol{v}\,d\boldsymbol{x} + \int_\Omega \mathrm{div}\,\boldsymbol{w}\,\mathrm{div}\,\boldsymbol{v}\,d\boldsymbol{x} = \int_\Omega \mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\,\mathrm{div}\,\boldsymbol{v}\,d\boldsymbol{x} \quad \forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega),$$

$$\int_\Omega (\boldsymbol{w} - \nabla r) \cdot \nabla q\,d\boldsymbol{x} = 0 \quad \forall q \in H_0^1(\Omega).$$

Then $\Delta r = \mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)$ and $\boldsymbol{w} = \nabla r$, and so $\|r\|_{H^2(\Omega)} + \|\boldsymbol{w}\|_{H^1(\Omega)} \leq C\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$. Taking $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}_h$, $q = p - p_h$ and using the error equations, we obtain

$$\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}^2 = \int_\Omega (\boldsymbol{w} - \boldsymbol{w}_I - \nabla[r - r_I]) \cdot (\boldsymbol{u} - \boldsymbol{u}_h - \nabla[p - p_h])\,d\boldsymbol{x}$$

$$+ \int_\Omega \mathrm{div}(\boldsymbol{w} - \boldsymbol{w}_I)\,\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\,d\boldsymbol{x}$$

for any $\boldsymbol{w}_I \in \boldsymbol{S}_h$ and $r_I \in W_h$. Taking $\boldsymbol{w}_I = \boldsymbol{\pi}_h \boldsymbol{w}$ and r_I a standard interpolant of r, the first integral on the right-hand side is bounded by

$$Ch\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}(\|p - p_h\|_{H^1(\Omega)} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}(\mathrm{div}, \Omega)})$$

$$\leq Ch^{r+2}\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}(\|p\|_{H^{r+1}(\Omega)} + \|\boldsymbol{u}\|_{H^{r+1}(\Omega)} + \|\mathrm{div}\,\boldsymbol{u}\|_{H^{r+1}(\Omega)}).$$

To bound the second integral, we note that, in the rectangular case, $\mathrm{div}\,\boldsymbol{\pi}_h \boldsymbol{w} = \Pi_h \,\mathrm{div}\,\boldsymbol{w}$ with Π_h the L^2-projection into $\mathrm{div}\,\boldsymbol{V}_h$, and also, in the rectangular case, $\mathrm{div}\,\boldsymbol{V}_h$ contains all piecewise polynomials of degree at most $r+1$, so $\|q - \Pi_h q\|_{L^2(\Omega)} \leq Ch^{r+2}\|q\|_{H^{r+2}(\Omega)}$ for all q. Therefore,

$$\int_\Omega \mathrm{div}(\boldsymbol{w} - \boldsymbol{w}_I)\,\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\,d\boldsymbol{x} = \int_\Omega \mathrm{div}\,\boldsymbol{w}[\mathrm{div}\,\boldsymbol{u} - \Pi_h(\mathrm{div}\,\boldsymbol{u})]\,d\boldsymbol{x}$$

$$\leq C\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}h^{r+2}\|\mathrm{div}\,\boldsymbol{u}\|_{H^{r+2}(\Omega)}.$$

Combining these estimates, we conclude that

$$\|\mathrm{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)} \leq Ch^{r+2}(\|p\|_{H^{r+1}(\Omega)} + \|\boldsymbol{u}\|_{H^{r+1}(\Omega)} + \|\mathrm{div}\,\boldsymbol{u}\|_{H^{r+2}(\Omega)}).$$

8. Numerical results. In this section, we illustrate our results with several numerical examples using two sequences of meshes. The first is a uniform mesh of the unit square into n^2 subsquares and the second is a mesh of trapezoids as shown in Figure 1(b) (with the notation of Theorem 3.1, here $\alpha = 1$ and $\beta = 1/3$). In the first of

TABLE 1

Errors and orders of convergence for the piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection into discontinuous $\boldsymbol{BDM}_1$ and discontinuous $\boldsymbol{BDFM}_2$.

Piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection into $\boldsymbol{BDM}_1$ on square meshes

n	$\|\boldsymbol{u} - \pi_h \boldsymbol{u}\|_{L^2(\Omega)}$			$\|\mathrm{div}(\boldsymbol{u} - \pi_h \boldsymbol{u})\|_{L^2(\Omega)}$		
	err.	%	order	err.	%	order
2	1.94e−02	13.010		2.11e−01	30.151	
4	5.08e−03	3.405	1.9	1.15e−01	16.428	0.9
8	1.28e−03	0.861	2.0	5.86e−02	8.375	1.0
16	3.22e−04	0.216	2.0	2.94e−02	4.207	1.0
32	8.05e−05	0.054	2.0	1.47e−01	2.106	1.0
64	2.01e−05	0.013	2.0	7.36e−03	1.053	1.0

Piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection into $\boldsymbol{BDM}_1$ on trapezoidal meshes

n	$\|\boldsymbol{u} - \pi_h \boldsymbol{u}\|_{L^2(\Omega)}$			$\|\mathrm{div}(\boldsymbol{u} - \pi_h \boldsymbol{u})\|_{L^2(\Omega)}$		
	err.	%	order	err.	%	order
2	2.57e−02	17.243		2.63e−01	37.646	
4	7.89e−03	5.291	1.7	1.83e−01	26.109	0.5
8	2.80e−03	1.879	1.5	1.50e−01	21.430	0.3
16	1.21e−03	0.811	1.2	1.40e−01	20.031	0.1
32	5.78e−04	0.387	1.1	1.37e−01	19.662	0.0
64	2.85e−04	0.191	1.0	1.37e−01	19.568	0.0

Piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection into $\boldsymbol{BDFM}_2$ on square meshes

n	$\|\boldsymbol{u} - \pi_h \boldsymbol{u}\|_{L^2(\Omega)}$			$\|\mathrm{div}(\boldsymbol{u} - \pi_h \boldsymbol{u})\|_{L^2(\Omega)}$		
	err.	%	order	err.	%	order
2	1.52e−02	10.206		5.27e−02	7.538	
4	3.80e−03	2.552	2.0	1.32e−02	1.884	2.0
8	9.51e−04	0.638	2.0	3.29e−03	0.471	2.0
16	2.38e−04	0.159	2.0	8.24e−04	0.118	2.0
32	5.94e−05	0.040	2.0	2.06e−04	0.029	2.0
64	1.49e−05	0.010	2.0	5.15e−05	0.007	2.0

Piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection into $\boldsymbol{BDFM}_2$ on trapezoidal meshes

n	$\|\boldsymbol{u} - \pi_h \boldsymbol{u}\|_{L^2(\Omega)}$			$\|\mathrm{div}(\boldsymbol{u} - \pi_h \boldsymbol{u})\|_{L^2(\Omega)}$		
	err.	%	order	err.	%	order
2	1.86e−02	12.502		6.85e−02	9.791	
4	5.07e−03	3.399	1.9	3.52e−02	5.040	1.0
8	1.38e−03	0.926	1.9	1.77e−02	2.538	1.0
16	4.29e−04	0.288	1.7	8.89e−03	1.271	1.0
32	1.66e−04	0.111	1.4	4.45e−03	0.636	1.0
64	7.56e−05	0.051	1.1	2.22e−03	0.318	1.0

these examples (see Table 1), we demonstrate the decreased orders of convergence of the $\boldsymbol{BDM}_1$ and $\boldsymbol{BDFM}_2$ spaces by computing the piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection of a simple smooth function, $\boldsymbol{u} = \mathrm{grad}[x_1(1 - x_1)x_2(1 - x_2)]$, into the discontinuous versions of these spaces. On a rectangular mesh, the space $\boldsymbol{BDFM}_2$ gives second order approximation of both components of the vector and of its divergence. This is

TABLE 2
Errors and orders of convergence for the mixed approximation to Poisson's equation.

$\boldsymbol{\mathcal{RT}}_0$ *on square meshes*

| n | $\|p - p_h\|_{L^2(\Omega)}$ | | | $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ | | | $\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ | | |
	err.	%	order	err.	%	order	err.	%	order
2	1.84e−02	55.28		6.09e−02	40.83		2.11e−01	30.15	
4	1.04e−02	31.07	0.8	3.32e−02	22.24	0.9	1.15e−01	16.43	0.9
8	5.33e−03	15.99	1.0	1.69e−02	11.34	1.0	5.86e−02	8.38	1.0
16	2.68e−03	8.05	1.0	8.49e−03	5.70	1.0	2.94e−02	4.21	1.0
32	1.34e−03	4.03	1.0	4.25e−03	2.85	1.0	1.47e−02	2.11	1.0

$\boldsymbol{\mathcal{RT}}_0$ *on trapezoidal meshes*

| n | $\|p - p_h\|_{L^2(\Omega)}$ | | | $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ | | | $\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ | | |
	err.	%	order	err.	%	order	err.	%	order
2	1.84e−02	55.08		6.34e−02	42.55		2.67e−01	38.14	
4	1.08e−02	32.37	0.8	3.63e−02	24.38	0.8	1.85e−01	26.51	0.5
8	5.60e−03	16.80	0.9	1.91e−02	12.83	0.9	1.53e−01	21.82	0.3
16	2.83e−03	8.48	1.0	9.81e−03	6.58	1.0	1.43e−01	20.42	0.1
32	1.42e−03	4.25	1.0	4.97e−03	3.33	1.0	1.40e−01	20.05	0.0

$\boldsymbol{\mathcal{ABF}}_0$ *on square meshes*

| n | $\|p - p_h\|_{L^2(\Omega)}$ | | | $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ | | | $\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ | | |
	err.	%	order	err.	%	order	err.	%	order
2	2.49e−02	74.59		6.89e−02	64.21		5.27e−02	7.54	
4	1.36e−02	40.65	0.9	3.42e−02	22.97	1.0	1.32e−02	1.88	2.0
8	7.03e−03	21.08	1.0	1.70e−02	11.43	1.0	3.29e−03	0.47	2.0
16	3.70e−03	11.10	0.9	8.51e−03	5.71	1.0	8.24e−04	0.12	2.0
32	1.93e−03	5.78	0.9	4.25e−03	2.85	1.0	2.06e−04	0.03	2.0

$\boldsymbol{\mathcal{ABF}}_0$ *on trapezoidal meshes*

| n | $\|p - p_h\|_{L^2(\Omega)}$ | | | $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ | | | $\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ | | |
	err.	%	order	err.	%	order	err.	%	order
2	2.31e−02	69.38		6.59e−02	44.20		6.91e−02	9.89	
4	1.33e−02	39.98	0.8	3.58e−02	24.04	0.9	3.58e−02	5.12	0.9
8	7.22e−03	21.66	0.9	1.85e−02	12.41	1.0	1.81e−02	2.58	1.0
16	3.84e−03	11.51	0.9	9.43e−03	6.33	1.0	9.05e−03	1.30	1.0
32	2.00e−03	5.99	0.9	4.77e−03	3.20	1.0	4.53e−03	0.65	1.0

confirmed in the approximation of the piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection. On a trapezoidal mesh, $\boldsymbol{\mathcal{BDFM}}_2$ gives only first order approximation of both components of the vector and of its divergence, and this is also confirmed in the approximation of the piecewise $\boldsymbol{H}(\mathrm{div}, \Omega)$ projection. On a rectangular mesh, the space $\boldsymbol{\mathcal{BDM}}_1$ gives second order approximation of both components of the vector, but only first order approximation of its divergence. On a trapezoidal mesh these orders of convergence are reduced to first order for the approximation of both components of the vector and the approximation of the divergence shows no convergence. These theoretical convergence orders are also confirmed in the computations. Although we do not in-

TABLE 3
Errors and orders of convergence for the least squares approximation to Poisson's equation.

$\mathcal{RT}_0$ *on square meshes*

n	$\|p - p_h\|_{L^2(\Omega)}$ err.	%	order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ err.	%	order	$\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ err.	%	order
2	2.61e−01	52.28		1.07e+00	48.03		5.78e+00	58.58	
4	7.71e−02	15.42	1.8	5.15e−01	23.19	1.1	3.09e+00	31.34	0.9
8	2.01e−02	4.01	1.9	2.53e−01	11.41	1.0	1.57e+00	15.94	1.0
16	5.07e−03	1.01	2.0	1.26e−01	5.68	1.0	7.90e−01	8.00	1.0
32	1.27e−03	0.25	2.0	6.30e−02	2.84	1.0	3.95e−01	4.01	1.0
64	3.18e−04	0.06	2.0	3.15e−02	1.42	1.0	1.98e−01	2.00	1.0

$\mathcal{RT}_0$ *on trapezoidal meshes*

n	$\|p - p_h\|_{L^2(\Omega)}$ err.	%	order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ err.	%	order	$\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ err.	%	order
2	2.95e−01	58.96		1.24e+00	55.74		6.03e+00	61.07	
4	1.08e−01	21.67	1.4	6.05e−01	27.26	1.0	3.68e+00	37.25	0.7
8	4.29e−02	8.58	1.3	3.10e−01	13.97	1.0	2.50e+00	25.37	0.6
16	2.51e−02	5.01	0.8	1.72e−01	7.74	0.9	2.09e+00	21.16	0.3
32	2.06e−02	4.12	0.3	1.13e−01	5.09	0.6	1.97e+00	19.96	0.1
64	1.95e−02	3.89	0.1	9.27e−02	4.17	0.3	1.94e+00	19.64	0.0

$\mathcal{ABF}_0$ *on square meshes*

n	$\|p - p_h\|_{L^2(\Omega)}$ err.	%	order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ err.	%	order	$\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ err.	%	order
2	1.42e−01	28.46		1.04e+00	46.77		2.19e+00	22.18	
4	3.35e−02	6.70	2.1	5.10e−01	22.98	1.0	5.88e−01	9.96	1.9
8	8.22e−03	1.64	2.0	2.53e−01	11.38	1.0	1.50e−01	1.52	2.0
16	2.04e−03	0.41	2.0	1.26e−01	5.67	1.0	3.76e−02	0.38	2.0
32	5.10e−04	0.10	2.0	6.30e−02	2.84	1.0	8.41e−03	0.10	2.0

$\mathcal{ABF}_0$ *on trapezoidal meshes*

n	$\|p - p_h\|_{L^2(\Omega)}$ err.	%	order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ err.	%	order	$\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega)}$ err.	%	order
2	1.89e−01	37.74		1.17e+00	52.86		3.0e+00	31.39	
4	5.49e−02	10.98	1.8	5.61e−01	25.24	1.1	1.12e+00	11.32	1.5
8	1.45e−02	2.89	1.9	2.80e−01	12.62	1.0	5.00e−01	5.07	1.2
16	3.67e−03	0.73	1.9	1.40e−01	6.32	1.0	2.42e−01	2.45	1.0
32	9.20e−04	0.18	2.0	7.02e−02	3.16	1.0	1.20e−01	1.21	1.0

clude the details of the computations, the same convergence orders are observed in computations of the $\boldsymbol{L}^2(\Omega)$, rather than the piecewise $\boldsymbol{H}(\operatorname{div}, \Omega)$ projection.

The second computation, reported in Table 2, illustrates our results on the convergence orders of $\mathcal{RT}_0$ and $\mathcal{ABF}_0$ for the approximation of Poisson's equation by the standard mixed finite element method. The exact solution is $p = x_1(1-x_1)x_2(1-x_2)$. As expected, on a trapezoidal mesh, $\mathcal{RT}_0$ gives a first order approximation to the scalar and vector variable (the same as on a rectangular mesh), but there is no con-

vergence of the approximation of the divergence of the vector variable in contrast to the standard first order approximation seen on rectangles. When $\boldsymbol{\mathcal{ABF}}_0$ is used instead, there is an improvement in the convergence order of the divergence of the vector variable.

Finally, Table 3 shows the difference in the convergence orders of $\boldsymbol{\mathcal{RT}}_0$ and $\boldsymbol{\mathcal{ABF}}_0$ coupled with $\mathcal{Q}_1$ for the scalar variable for the approximation of Poisson's equation by a standard least squares finite element method. Again the exact solution is $p = x_1(1 - x_1)x_2(1 - x_2)$. When $\boldsymbol{\mathcal{RT}}_0$ is used, the poor approximation of the divergence on trapezoidal meshes results in poor approximation of both the scalar and vector variable, while on a rectangle the scalar variable is approximated to second order and the vector variable and its divergence to first order. When $\boldsymbol{\mathcal{ABF}}_0$ is used instead, one achieves second order convergence for the scalar variable and first order convergence for the vector variable on both rectangular and quadrilateral meshes. The divergence of the vector variable is approximated to second order on rectangles and to first order on trapezoids, as predicted by the theory.

REFERENCES

[1] D. N. Arnold, D. Boffi, and R. S. Falk, *Approximation by quadrilateral finite elements*, Math. Comp., 71 (2002), pp. 909–922.

[2] S. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[3] F. Brezzi, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge, 8 (1974), pp. 129–151.

[4] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[5] J. Douglas Jr. and J. E. Roberts, *Global estimates for mixed methods for second order elliptic equations*, Math. Comp., 44 (1985), pp. 39–52.

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

[7] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier–Stokes Equations*, Springer-Verlag, Berlin, 1986.

[8] Y. Kuznetsov and S. Repin, *New mixed finite element method on polygonal and polyhedral meshes*, Russian J. Numer. Anal. Math. Modelling, 18 (2003), pp. 261–278.

[9] R. L. Naff, T. F. Russel, and J. D. Wilson, *Shape functions for velocity interpolation in general hexahedral cells*, Comput. Geosci., 6 (2002), pp. 285–314.

[10] A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1368–1377.

[11] J. Shen, *Mixed Finite Element Methods on Distorted Rectangular Grids*, Technical report, Institute for Scientific Computation, Texas A&M University, College Station, TX, 1994.

[12] J. M. Thomas, *Sur l'analyse numerique des methodes d'elements finis hybrides et mixtes*, Ph.D. Thesis, Université Pierre et Marie Curie, Paris, France, 1977.

[13] Z. Zhang, *Analysis of some quadrilateral nonconforming elements for incompressible elasticity*, SIAM J. Numer. Anal., 34 (1997), pp. 640–663.