World Scientific
www.worldscientific.com

Higher order multipoint flux mixed finite element methods on quadrilaterals and hexahedra

Ilona Ambartsumyan* and Eldar Khattatov†

Oden Institute for Computational Engineering and Sciences,
The University of Texas at Austin,
201 East 24th St, Austin, TX 78712, USA
*ailona@austin.utexas.edu
†ekhattatov@austin.utexas.edu

Jeonghun J. Lee

Department of Mathematics, Baylor University,
One Bear Place #97328, Waco, TX 76798, USA
jeonghun_lee@baylor.edu

Ivan Yotov‡

Department of Mathematics, University of Pittsburgh,
301 Thackeray Hall, Pittsburgh, PA 15260, USA
yotov@math.pitt.edu

We develop higher order multipoint flux mixed finite element (MFMFE) methods for solving elliptic problems on quadrilateral and hexahedral grids that reduce to cell-based pressure systems. The methods are based on a new family of mixed finite elements, which are enhanced Raviart–Thomas spaces with bubbles that are curls of specially chosen polynomials. The velocity degrees of freedom of the new spaces can be associated with the points of tensor-product Gauss–Lobatto quadrature rules, which allows for local velocity elimination and leads to a symmetric and positive definite cell-based system for the pressures. We prove optimal kth order convergence for the velocity and pressure in their natural norms, as well as $(k + 1)$st order superconvergence for the pressure at the Gauss points. Moreover, local postprocessing gives a pressure that is superconvergent of order $(k + 1)$ in the full L^2-norm. Numerical results illustrating the validity of our theoretical results are included.

Keywords: Mixed finite element methods; multipoint flux, high order methods; finite volume schemes.

AMS Subject Classification: 65N08, 65N12, 65N15, 65N30, 76S05

‡Corresponding author

1. Introduction

Mixed finite element (MFE) methods[16,37] are commonly used for modeling of fluid flow and transport, as they provide accurate and locally mass conservative velocities and robustness with respect to heterogeneous, anisotropic, and discontinuous coefficients. A disadvantage of the MFE methods in their standard form is that they result in coupled velocity-pressure algebraic systems of saddle-point type, which restricts the use of efficient iterative solvers. To address this issue, there has been extensive work on developing modifications of MFE methods that can be reduced to positive definite systems, such as hybridization[10,16] or relating them to cell-centered finite difference or finite volume methods. In the latter approach, a common technique is to employ special quadrature rules, also referred to as mass lumping, that allow for local velocity elimination, resulting in cell-centered pressure systems. Early works of Refs. 13, 38 and 44 based on the lowest order Raviart–Thomas (RT_0) spaces[36] were limited to two-point flux approximations, which were not robust for general quadrilateral grids or tensor-valued coefficients. An extension to higher order RT spaces, as well as the second-order Brezzi–Douglas–Fortin–Marini ($BDFM_2$) spaces[16] was developed in Ref. 18, but was also limited to rectangular grids and diagonal tensor coefficients. The expanded MFE method[6,7] was designed to handle full tensor coefficients and general grids, but suffered from reduced convergence for problems with discontinuous coefficients.

More recently, a special MFE method, the multipoint flux mixed finite element (MFMFE) method[27,47] was developed, which reduces to cell-centered finite differences on quadrilateral, hexahedral and simplicial grids, and exhibits robust performance for discontinuous full tensor coefficients. The method was motivated by the multipoint flux approximation (MPFA) method,[1–3,21,22] which was developed as a finite volume method. Unlike the MPFA method, the variational formulation of the MFMFE method allows for its complete theoretical study of well-posedness and convergence. The MFMFE method is based on the lowest order Brezzi–Douglas–Marini (BDM_1) space[15,35] on simplices and quadrilaterals, and an enhanced Brezzi–Douglas–Duran–Fortin ($BDDF_1$) space[14,27] on hexahedra. The method utilizes the trapezoidal quadrature rule for the velocity mass matrix, which reduces it to a block-diagonal form with blocks associated with mesh vertices. The velocities can then be easily eliminated, resulting in a cell-centered pressure system. A similar approach was also presented independently in Ref. 17 for simplicial grids, and a related formulation based on a broken Raviart–Thomas space was developed in works of Refs. 29 and 30. Motivated by the work in Ref. 30, a nonsymmetric version of the MFMFE method designed to converge on general quadrilateral and hexahedral grids was developed in Ref. 45. A multiscale mortar MFMFE method on multiple subdomains with non-matching grids was proposed in Ref. 46. In the work of Ref. 33, a local flux mimetic finite difference method was developed on polyhedral grids, exploring connections to the MFMFE and MPFA methods, see also related work in Refs. 28 and 40. Furthermore, on simplicial grids and for problems

with full tensor coefficients, using the MPFA principle, it was shown in Refs. 41, 42 and 48 that the RT_0 MFE method can be related to a finite volume method with one pressure unknown per element.

To the best of our knowledge, the aforementioned MPFA and MFMFE methods with theoretical convergence proofs are limited to the lowest order approximation. In this paper, we develop a family of arbitrary order symmetric MFMFE methods on h^2-perturbed quadrilateral and hexahedral grids. The main obstacle in extending the original lowest order BDM_1 and $BDDF_1$ MFMFE methods to higher order is that the degrees of freedom of their higher order versions cannot be associated with tensor-product quadrature rules. To circumvent this difficulty, we construct a new family of mixed finite elements fulfilling this requirement. A key of the construction is the finite element exterior calculus framework,[11,12] which is used in the extension of MFMFE to Hodge Laplace equations.[31] However, we consider only the two- and three-dimensional cases with $H(\mathrm{div})$ element in this paper, so no prerequisite of the exterior calculus language is necessary. The new spaces are enhanced Raviart–Thomas spaces with bubbles that are curls of specially chosen polynomials, so that each component of the velocity vector is of dimension $\mathcal{Q}^k(\mathbb{R}^d)$ and the velocity degrees of freedom can be associated with the points of a tensor-product Gauss–Lobatto quadrature rule.[4] The application of this quadrature rule leads to a block-diagonal velocity mass matrix with blocks corresponding to the nodes associated with the velocity degrees of freedom. This allows for a local elimination of the fluxes in terms of the pressures from the surrounding elements, either sharing a vertex, or an edge/face. This procedure results in a symmetric and positive-definite cell-based system for the pressures with a compact stencil, allowing for efficient solvers to be used. The proposed technique allows for more straightforward and efficient implementation and results in reduced computational time. We remark that the lowest order version of our new elements has the same number of degrees of freedom as the elements used in previous MFMFE methods, but they are *different* elements. This work is *not* a direct extension of the previous MFMFE methods to higher order, but a new framework for explicit construction of higher order MFMFE methods.

We present well-posedness and convergence analysis of the proposed family of higher order methods. To this end, we establish unisolvency and approximation properties of arbitrary order k of the new family of enhanced Raviart–Thomas family of spaces. Since we study the symmetric version of the MFMFE method, which relies on mapping to a reference element via the Piola transformation, the analysis is limited to h^2-perturbed parallelograms or parallelepipeds, similar to the restriction in the lowest order symmetric MFMFE method.[27,47] The convergence analysis combines MFE analysis tools with quadrature error analysis, using that the Gauss–Lobatto quadrature rule possesses sufficient accuracy to preserve the order of convergence. We establish convergence of kth order for the velocity in the $H(\mathrm{div})$-norm and the pressure in the L^2-norm. We also employ a duality argument to show that the numerical pressure is $(k+1)$st order superconvergent to the L^2-projection

of the pressure in the finite element space, which implies superconvergence at the Gauss points. Moreover, we show that a variant of the local postprocessing developed in Ref. 39 results in a pressure that is $(k + 1)$st order accurate in the full L^2-norm. All theoretical results are verified numerically.

The rest of the paper is organized as follows. The new family of finite element spaces and the general order MFMFE methods are developed in Sec. 2. The error analyses for the velocity and pressure are presented in Secs. 3 and 4, respectively. Numerical experiments are presented in Sec. 5.

2. Definition of the Method

2.1. *Preliminaries*

We consider a second-order elliptic PDE written as a system of two first-order equations,

$$\mathbf{u} = -K\nabla p, \quad \nabla \cdot \mathbf{u} = f \text{ in } \Omega, \tag{2.1}$$

$$p = g \text{ on } \Gamma_D, \quad \mathbf{u} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N, \tag{2.2}$$

where $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ is an open bounded polytopal domain with a boundary $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ such that $\Gamma_D \cap \Gamma_N = \emptyset$, with measure$(\Gamma_D) > 0$. Here, $\mathbf{n}$ is the outward unit normal vector field on $\partial\Omega$, and K is symmetric and uniformly positive definite tensor satisfying, for some $0 < k_0 < k_1 < \infty$,

$$k_0\xi^T\xi \leq \xi^T K(\mathbf{x})\xi \leq k_1\xi^T\xi, \quad \forall \mathbf{x} \in \Omega, \ \forall \xi \in \mathbb{R}^d. \tag{2.3}$$

In applications related to modeling flow in porous media, p is the pressure, $\mathbf{u}$ is the Darcy velocity, and K represents the permeability tensor divided by the viscosity. The above choice of boundary conditions is made for the sake of simplicity. More general boundary conditions, including nonhomogeneous full Neumann ones, can also be treated.

Throughout the paper, we will use the following standard notation. For a domain $G \subset \mathbb{R}^d$, the $L^2(G)$ inner product and norm for scalar- and vector-valued functions are denoted by $(\cdot, \cdot)_G$ and $\|\cdot\|_G$, respectively. The norms and seminorms of the Sobolev spaces $W^{s,p}(G)$, $s \in \mathbb{R}$, $p \geq 1$ are denoted by $\|\cdot\|_{s,p,G}$ and $|\cdot|_{s,p,G}$, respectively. Conventionally, the norms and seminorms of Hilbert spaces $H^s(G)$ are denoted by $\|\cdot\|_{s,G}$ and $|\cdot|_{s,G}$, respectively. We omit G in the subscript if $G = \Omega$. For a section of the domain or element boundary $S \subset \mathbb{R}^{d-1}$ we write $\langle\cdot, \cdot\rangle_S$ and $\|\cdot\|_S$ for the $L^2(S)$ inner product (or duality pairing) and norm, respectively. For a tensor-valued function M, let $\|M\|_\alpha = \max_{i,j}\|M_{i,j}\|_\alpha$ for any norm $\|M\|_\alpha$. We will also use the space

$$H(\text{div}; \Omega) = \{\mathbf{v} \in L^2(\Omega, \mathbb{R}^d) : \nabla \cdot \mathbf{v} \in L^2(\Omega)\}$$

equipped with the norm

$$\|\mathbf{v}\|_{\text{div}} = (\|\mathbf{v}\|^2 + \|\nabla \cdot \mathbf{v}\|^2)^{1/2}.$$

The weak formulation for (2.1)–(2.2) reads as follows: find $(\mathbf{u}, p) \in \mathbf{V} \times W$ such that

$$\left(K^{-1}\mathbf{u}, \mathbf{v}\right) - (p, \nabla \cdot \mathbf{v}) = -\langle g, \mathbf{v} \cdot \mathbf{n} \rangle_{\Gamma_D}, \quad \mathbf{v} \in \mathbf{V}, \tag{2.4}$$

$$\left(\nabla \cdot \mathbf{u}, w\right) = (f, w), \quad w \in W, \tag{2.5}$$

where

$$\mathbf{V} = \{\mathbf{v} \in H(\operatorname{div}; \Omega) : \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N\}, \quad W = L^2(\Omega).$$

It was shown[16,37] that (2.4)–(2.5) has a unique solution.

2.2. *A finite element mapping*

Let $\mathcal{T}_h$ be a finite element partition of Ω consisting of quadrilaterals in 2d or hexahedra in 3d, where $h = \max_{E \in \mathcal{T}_h} \operatorname{diam}(E)$. We assume $\mathcal{T}_h$ to be shape regular and quasi-uniform.[22] For any element $E \in \mathcal{T}_h$ there exists a bilinear (trilinear) bijection mapping $F_E : \hat{E} \to E$, where $\hat{E} = [-1, 1]^d$ is the reference square (cube). Denote the Jacobian matrix by DF_E, and let $J_E = |\det(DF_E)|$. Denote the inverse mapping by F_E^{-1}, its Jacobian matrix by DF_E^{-1}, and let $J_{F_E^{-1}} = |\det(DF_E^{-1})|$. For $\hat{\mathbf{x}} = F_E^{-1}(\mathbf{x})$ we have that

$$DF_E^{-1}(\mathbf{x}) = (DF_E)^{-1}(\hat{\mathbf{x}}), \quad J_{F_E^{-1}}(\mathbf{x}) = \frac{1}{J_E(\hat{\mathbf{x}})}.$$

Denote by $\hat{\mathbf{r}}_i$, $i = 1, \ldots, 2^d$, the vertices of $\hat{E}$, where $\hat{\mathbf{r}}_1 = (0, 0)^T$, $\hat{\mathbf{r}}_2 = (1, 0)^T$, $\hat{\mathbf{r}}_3 = (1, 1)^T$, and $\hat{\mathbf{r}}_4 = (0, 1)^T$ in 2d, and $\hat{\mathbf{r}}_1 = (0, 0, 0)^T$, $\hat{\mathbf{r}}_2 = (1, 0, 0)^T$, $\hat{\mathbf{r}}_3 = (1, 1, 0)^T$, $\hat{\mathbf{r}}_4 = (0, 1, 0)^T$, $\hat{\mathbf{r}}_5 = (0, 0, 1)^T$, $\hat{\mathbf{r}}_6 = (1, 0, 1)^T$, $\hat{\mathbf{r}}_7 = (1, 1, 1)^T$, and $\hat{\mathbf{r}}_8 = (0, 1, 1)^T$ in 3d. Let $\mathbf{r}_i$, $i = 1, \ldots, 2^d$, be the corresponding vertices of element E. The outward unit normal vector fields to the facets of E and $\hat{E}$ are denoted by $\mathbf{n}_i$ and $\hat{\mathbf{n}}_i$, $i = 1, \ldots, 2d$, respectively, where facet is a face in 3d or an edge in 2d. The bilinear (trilinear) mapping is given by

$$F_E(\hat{\mathbf{r}}) = \mathbf{r}_1 + \mathbf{r}_{21}\hat{x} + \mathbf{r}_{41}\hat{y} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x}\hat{y}, \quad \text{in 2d,} \tag{2.6}$$

$$F_E(\hat{\mathbf{r}}) = \mathbf{r}_1 + \mathbf{r}_{21}\hat{x} + \mathbf{r}_{41}\hat{y} + \mathbf{r}_{51}\hat{z} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x}\hat{y} + (\mathbf{r}_{65} - \mathbf{r}_{21})\hat{x}\hat{z}$$

$$+ (\mathbf{r}_{85} - \mathbf{r}_{41})\hat{y}\hat{z} + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{x}\hat{y}\hat{z}, \quad \text{in 3d,} \tag{2.7}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. For the 3d case, we note that the elements can have nonplanar faces.

Let $\hat{\phi}(\hat{\mathbf{x}})$ be defined on $\hat{E}$, and let $\phi = \hat{\phi} \circ F_E^{-1}$. Using the classical formula $\nabla \phi = (DF_E^{-1})^T \hat{\nabla} \hat{\phi}$, it is easy to see that for any facet $e_i \subset \partial E$

$$\mathbf{n}_i = \frac{1}{J_{e_i}} J_E (DF_E^{-1})^T \hat{\mathbf{n}}_i, \quad J_{e_i} = |J_E (DF_E^{-1})^T \hat{\mathbf{n}}_i|_{\mathbb{R}^d}, \tag{2.8}$$

where $|\cdot|_{\mathbb{R}^d}$ denotes the Euclidean vector norm in $\mathbb{R}^d$. Another straightforward calculation shows that, for all element types, the mapping definitions and the shape-regularity and quasi-uniformity of the grids imply that

$$\|DF_E\|_{0,\infty,\hat{E}} \sim h, \quad \|J_E\|_{0,\infty,\hat{E}} \sim h^d, \quad \text{and}$$
$$\|DF_E^{-1}\|_{0,\infty,E} \sim h^{-1}, \quad \|J_{F_E^{-1}}\|_{0,\infty,E} \sim h^{-d}, \tag{2.9}$$

where the notation $a \sim b$ means that there exist positive constants c_0, c_1 independent of h such that $c_0 b \le a \le c_1 b$.

2.3. The Raviart–Thomas mixed finite element spaces

Let $\mathcal{P}^k$ denote the space of polynomials of total degree $\le k$ and let $\mathcal{Q}^k$ denote the space of polynomials of degree $\le k$ in each variable. We will make use of the Raviart–Thomas spaces for the construction of the spaces needed for the proposed method. The RT_k spaces are defined for $k \ge 0$ on the reference cube as

$$\hat{\mathbf{V}}_{RT}^k(\hat{E}) = \begin{pmatrix} \mathcal{Q}^k + \mathcal{Q}^k \hat{x} \\ \mathcal{Q}^k + \mathcal{Q}^k \hat{y} \\ \mathcal{Q}^k + \mathcal{Q}^k \hat{z} \end{pmatrix}, \quad \hat{W}^k(\hat{E}) = \mathcal{Q}^k(\hat{E}). \tag{2.10}$$

The definition on the reference square can be obtained naturally from the one above. It holds that

$$\hat{\nabla} \cdot \hat{\mathbf{V}}_{RT}^k(\hat{E}) = \hat{W}^k(\hat{E}) \quad \text{and} \quad \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}} \in \mathcal{Q}^k(\hat{e}) \quad \forall \hat{\mathbf{v}} \in \hat{\mathbf{V}}_{RT}^k(\hat{E}), \ \forall \hat{e} \subset \partial \hat{E}. \tag{2.11}$$

The projection operator $\hat{\Pi}_{RT}^k : H^1(\hat{E}, \mathbb{R}^d) \to \hat{\mathbf{V}}_{RT}^k(\hat{E})$ satisfies

$$\text{for } k \ge 0, \quad \langle (\hat{\mathbf{q}} - \hat{\Pi}_{RT}^k \hat{\mathbf{q}}) \cdot \mathbf{n}_{\hat{e}}, \hat{p} \rangle_{\hat{e}} = 0, \quad \forall \hat{p} \in \mathcal{Q}^k(\hat{e}), \ \forall \hat{e} \subset \partial \hat{E}, \tag{2.12}$$

$$\text{for } k \ge 1, \quad \left(\hat{\Pi}_{RT}^k \hat{\mathbf{q}} - \hat{\mathbf{q}}, \hat{\mathbf{p}} \right)_{\hat{E}} = 0, \quad \forall \hat{\mathbf{p}} \in \begin{cases} \begin{pmatrix} \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{P}^k(\hat{y}) \\ \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{P}^k(\hat{x}) \end{pmatrix} & \text{in 2d}, \\ \begin{pmatrix} \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z}) \\ \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z}) \\ \mathcal{P}^{k-1}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y}) \end{pmatrix} & \text{in 3d}. \end{cases} \tag{2.13}$$

The Raviart–Thomas spaces on any quadrilateral or hexahedral element $E \in \mathcal{T}_h$ are defined via the transformations

$$\mathbf{v} \leftrightarrow \hat{\mathbf{v}} : \mathbf{v} = \frac{1}{J_E} DF_E \hat{\mathbf{v}} \circ F_E^{-1}, \quad w \leftrightarrow \hat{w} : w = \hat{w} \circ F_E^{-1}, \tag{2.14}$$

where the contravariant Piola transformation is used for the velocity space. Under this transformation, the normal components of the velocity vectors on the facets

are preserved. In particular,[16]

$$\forall \, \hat{\mathbf{v}} \in \hat{\mathbf{V}}_{RT}^k(\hat{E}), \quad \forall \, \hat{w} \in \hat{W}^k(\hat{E}),$$

$$\left(\nabla \cdot \mathbf{v}, w\right)_E = \left(\hat{\nabla} \cdot \hat{\mathbf{v}}, \hat{w}\right)_{\hat{E}} \quad \text{and} \quad \langle \mathbf{v} \cdot \mathbf{n}_e, w \rangle_e = \langle \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}}, \hat{w} \rangle_{\hat{e}}, \tag{2.15}$$

which imply

$$\mathbf{v} \cdot \mathbf{n}_e = \frac{1}{J_e} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}}, \quad \nabla \cdot \mathbf{v}(\mathbf{x}) = \left(\frac{1}{J_E} \hat{\nabla} \cdot \hat{\mathbf{v}}\right) \circ F_E^{-1}(\mathbf{x}). \tag{2.16}$$

The RT_k spaces on $\mathcal{T}_h$ are given by

$$\mathbf{V}_{RT,h}^k = \left\{ \mathbf{v} \in \mathbf{V} : \mathbf{v}|_E \leftrightarrow \hat{\mathbf{v}}, \, \hat{\mathbf{v}} \in \hat{\mathbf{V}}_{RT}^k(\hat{E}), \, E \in \mathcal{T}_h \right\},$$

$$W_h^k = \left\{ w \in W : w|_E \leftrightarrow \hat{w}, \, \hat{w} \in \hat{W}^k(\hat{E}), \, E \in \mathcal{T}_h \right\}. \tag{2.17}$$

Using the Piola transformation, we define a projection operator Π_{RT}^k from $\mathbf{V} \cap H^1(\Omega, \mathbb{R}^d)$ onto $\mathbf{V}_{RT,h}^k$ satisfying on each element

$$\Pi_{RT}^k \mathbf{q} \leftrightarrow \widehat{\Pi_{RT}^k \mathbf{q}}, \quad \widehat{\Pi_{RT}^k \mathbf{q}} = \hat{\Pi}_{RT}^k \hat{\mathbf{q}}. \tag{2.18}$$

Using (2.16), (2.12)–(2.13) and (2.18), it is straightforward to show that $\Pi_{RT}^k \mathbf{q} \cdot \mathbf{n}$ is continuous across element facets, so $\Pi_{RT}^k \mathbf{q} \in H(\mathrm{div}; \Omega)$. Similarly, one can see that $\Pi_{RT}^k \mathbf{q} \cdot \mathbf{n} = 0$ on Γ_N if $\mathbf{q} \cdot \mathbf{n} = 0$ on Γ_N, so $\Pi_{RT}^k \mathbf{q} \in \mathbf{V}_{RT,h}^k$. Details of these arguments can be found in Refs. 9, 16, 27, 43, 47.

2.4. *Enhanced Raviart–Thomas finite elements*

In this section we develop a new family of enhanced Raviart–Thomas spaces, which is used in our method. We present the definitions of shape functions and degrees of freedom and discuss their unisolvency. The idea of the construction is to enhance the Raviart–Thomas spaces with bubbles that are curls of specially chosen polynomials, so that each component of the velocity vector is of dimension $Q^k(\mathbb{R}^d)$ and the velocity degrees of freedom can be associated with the points of a tensor-product Gauss–Lobatto quadrature rule.

2.4.1. *Shape functions*

In this subsection we adopt a convention for compact notation that $w^{-1} = 0$ for a polynomial variable w unless it is multiplied by w. For example, it holds that $\hat{x}^{-1}(\hat{x}, \hat{y}, \hat{x}^2 \hat{z})^T = (1, 0, \hat{x}\hat{z})^T$. For $k \geq 1$ and integers d_1, d_2, d_3, define

$$\mathcal{B}_1^k(\hat{E}) = \mathrm{span} \left\{ \begin{pmatrix} \hat{x}^{d_1} \hat{y}^{d_2} \hat{z}^{d_3} \\ 0 \\ 0 \end{pmatrix} : 0 \leq d_1, d_2, d_3 \leq k, d_2 = k \text{ or } d_3 = k \right\},$$

$$\boldsymbol{\mathcal{B}}_2^k(\hat{E}) = \text{span}\left\{ \begin{pmatrix} 0 \\ \hat{x}^{d_1}\hat{y}^{d_2}\hat{z}^{d_3} \\ 0 \end{pmatrix} : 0 \le d_1, d_2, d_3 \le k, d_1 = k \text{ or } d_3 = k \right\},$$

$$\boldsymbol{\mathcal{B}}_3^k(\hat{E}) = \text{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ \hat{x}^{d_1}\hat{y}^{d_2}\hat{z}^{d_3} \end{pmatrix} : 0 \le d_1, d_2, d_3 \le k, d_1 = k \text{ or } d_2 = k \right\}$$

on the reference element $\hat{E}$. While the above construction was done explicitly in 3d, it translates naturally to 2d by omitting the $\hat{z}$ terms. We now define the space $\boldsymbol{\mathcal{B}}^k$ as

$$\boldsymbol{\mathcal{B}}^k(\hat{E}) = \bigcup_{i=1}^d \boldsymbol{\mathcal{B}}_i^k. \tag{2.19}$$

It is clear from the above definition that $\mathcal{Q}^k(\hat{E}, \mathbb{R}^d) = \hat{\mathbf{V}}_{RT}^{k-1}(\hat{E}) \oplus \boldsymbol{\mathcal{B}}^k(\hat{E})$ in both 2d and 3d.

For $\hat{\mathbf{q}} \in \boldsymbol{\mathcal{B}}^k(\hat{E})$, we then consider $\hat{\nabla} \times (\hat{\mathbf{x}} \times \hat{\mathbf{q}})$. Here, we use the regular curl and cross product operators in 3d. The cross product applies to a 2d vector by representing the vector as a 3d one, with zeroed out third component, resulting in a scalar function, i.e. $\hat{\mathbf{x}} \times \hat{\mathbf{q}} = \hat{x}q_2 - \hat{y}q_1$ for $\hat{\mathbf{q}} = (q_1, q_2)^T$. In 2d, $\hat{\nabla} \times$ applies to a scalar function ϕ by representing the scalar function as a 3d vector with zero first and second components, and the first and second components of the result is defined as $\hat{\nabla} \times \phi$, i.e. $\hat{\nabla} \times \phi = (-\partial_2 \phi, \partial_1 \phi)^T$. Therefore, if $\hat{\mathbf{q}} = (q_1, 0)^T$ with $q_1 = \hat{x}^{a_1}\hat{y}^{a_2}$,

$$\hat{\nabla} \times (\hat{\mathbf{x}} \times \hat{\mathbf{q}}) = \hat{x}^{a_1-1}\hat{y}^{a_2} \begin{pmatrix} (a_2 + 1)\hat{x} \\ -a_1\hat{y} \end{pmatrix}.$$

We are now ready to construct a space isomorphic to $\boldsymbol{\mathcal{B}}^k(\hat{E})$, which is better suited for the analysis as well as for practical implementation. More precisely, we define

$$\tilde{\boldsymbol{\mathcal{B}}}_i^k(\hat{E}) = \hat{\nabla} \times (\hat{\mathbf{x}} \times \boldsymbol{\mathcal{B}}_i^k(\hat{E})), \quad i = 1, \dots, d, \quad \text{and} \quad \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}) = \bigcup_{i=1}^d \tilde{\boldsymbol{\mathcal{B}}}_i^k(\hat{E}).$$

One can check that in 2d,

$$\tilde{\boldsymbol{\mathcal{B}}}_1^k(\hat{E}) = \text{span}\left\{ \hat{x}^{a_1-1}\hat{y}^{a_2} \begin{pmatrix} (a_2 + 1)\hat{x} \\ -a_1\hat{y} \end{pmatrix} : 0 \le a_1, a_2 \le k, a_2 = k \right\}, \tag{2.20}$$

$$\tilde{\boldsymbol{\mathcal{B}}}_2^k(\hat{E}) = \text{span}\left\{ \hat{x}^{b_1}\hat{y}^{b_2-1} \begin{pmatrix} -b_2\hat{x} \\ (b_1 + 1)\hat{y} \end{pmatrix} : 0 \le b_1, b_2 \le k, b_1 = k \right\}, \tag{2.21}$$

and in 3d,

$$\tilde{\boldsymbol{\mathcal{B}}}_1^k(\hat{E}) = \text{span}\left\{ \hat{x}^{a_1-1}\hat{y}^{a_2}\hat{z}^{a_3} \begin{pmatrix} (a_2 + a_3 + 2)\hat{x} \\ -a_1\hat{y} \\ -a_1\hat{z} \end{pmatrix} : \begin{array}{l} 0 \le a_1, a_2, a_3 \le k, \\ a_2 = k \text{ or } a_3 = k \end{array} \right\}, \tag{2.22}$$

$$\tilde{\boldsymbol{\mathcal{B}}}_2^k(\hat{E}) = \text{span}\left\{ \hat{x}^{b_1}\hat{y}^{b_2-1}\hat{z}^{b_3} \begin{pmatrix} -b_2\hat{x} \\ (b_1 + b_3 + 2)\hat{y} \\ -b_2\hat{z} \end{pmatrix} : \begin{array}{l} 0 \le b_1, b_2, b_3 \le k, \\ b_1 = k \text{ or } b_3 = k \end{array} \right\}, \quad (2.23)$$

$$\tilde{\boldsymbol{\mathcal{B}}}_3^k(\hat{E}) = \text{span}\left\{ \hat{x}^{c_1}\hat{y}^{c_2}\hat{z}^{c_3-1} \begin{pmatrix} -c_3\hat{x} \\ -c_3\hat{y} \\ (c_1 + c_2 + 2)\hat{z} \end{pmatrix} : \begin{array}{l} 0 \le c_1, c_2, c_3 \le k, \\ c_1 = k \text{ or } c_2 = k \end{array} \right\}. \quad (2.24)$$

We define the enhanced Raviart–Thomas space $\hat{\mathbf{V}}^k(\hat{E})$ as

$$\hat{\mathbf{V}}^k(\hat{E}) = \hat{\mathbf{V}}_{RT}^{k-1}(\hat{E}) + \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}). \quad (2.25)$$

Theorem 2.1. *The sum (2.25) is a direct sum, i.e.* $\hat{\mathbf{V}}^k(\hat{E}) = \hat{\mathbf{V}}_{RT}^{k-1}(\hat{E}) \oplus \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$, *and* $\dim \hat{\mathbf{V}}^k(\hat{E}) = \dim \mathcal{Q}^k(\hat{E}, \mathbb{R}^d)$.

Proof. We will prove that the space $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$ is isomorphic to $\boldsymbol{\mathcal{B}}^k(\hat{E})$. It suffices to show that the map $\hat{\mathbf{q}} \mapsto \hat{\nabla} \times (\hat{\mathbf{x}} \times \hat{\mathbf{q}})$ is injective on $\boldsymbol{\mathcal{B}}^k(\hat{E})$. To see it, suppose that a linear combination of the elements of (2.22)–(2.24) is zero. Note that all elements in each space of (2.22)–(2.24) have distinct polynomials degrees. Therefore, for a component of fixed degrees of $\hat{x}$, $\hat{y}$, $\hat{z}$ in the linear combination, only one element of each space is used to generate the component. This implies that

$$\alpha\hat{x}^{a_1-1}\hat{y}^{a_2}\hat{z}^{a_3} \begin{pmatrix} (a_2 + a_3 + 2)\hat{x} \\ -a_1\hat{y} \\ -a_1\hat{z} \end{pmatrix} + \beta\hat{x}^{b_1}\hat{y}^{b_2-1}\hat{z}^{b_3} \begin{pmatrix} -b_2\hat{x} \\ (b_1 + b_3 + 2)\hat{y} \\ -b_2\hat{z} \end{pmatrix}$$

$$+ \gamma\hat{x}^{c_1}\hat{y}^{c_2}\hat{z}^{c_3-1} \begin{pmatrix} -c_3\hat{x} \\ -c_3\hat{y} \\ (c_1 + c_2 + 2)\hat{z} \end{pmatrix} = 0,$$

with some coefficients α, β, γ and

$$a_1 = b_1 + 1 = c_1 + 1, \quad b_2 = a_2 + 1 = c_2 + 1, \quad c_3 = a_3 + 1 = b_3 + 1. \quad (2.26)$$

We will prove that $\alpha = \beta = \gamma = 0$. If $a_2 = k$, then $\beta = 0$ due to $0 \le a_i, b_i, c_i \le k$ and (2.26). Comparing the components of the above equation, we have

$$-\alpha a_1 - \gamma(a_3 + 1) = 0, \quad -\alpha a_1 + \gamma(a_1 + a_2 + 1) = 0,$$

and therefore $\alpha = \gamma = 0$. Similarly, $\gamma = 0$ if $a_3 = k$ due to (2.26), and a similar argument gives

$$-\alpha a_1 - \beta(a_3 + 1) = 0, \quad -\alpha a_1 + \beta(a_1 + a_2 + 1) = 0,$$

which results in $\alpha = \beta = 0$. Since this argument holds for any component of the same polynomial degrees, the map $\hat{\mathbf{q}} \mapsto \hat{\nabla} \times (\hat{\mathbf{x}} \times \hat{\mathbf{q}})$ is injective on $\boldsymbol{\mathcal{B}}^k(\hat{E})$,

and therefore it is an isomorphism from $\boldsymbol{\mathcal{B}}^k(\hat{E})$ to $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$. This implies that the monomials in (2.20)–(2.21) and (2.22)–(2.24) form a basis of $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$. Note that every element of $\tilde{\boldsymbol{\mathcal{B}}}_i^k(\hat{E})$ in (2.20)–(2.24) contains at least one entry with a variable of degree $k+1$, therefore $\hat{\mathbf{V}}_{RT}^{k-1}(\hat{E}) \cap \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}) = \{0\}$, i.e. the sum (2.25) is a direct sum. This implies that $\dim \hat{\mathbf{V}}^k(\hat{E}) = \dim \mathcal{Q}^k(\hat{E}, \mathbb{R}^d)$.   □

2.4.2. Degrees of freedoms and unisolvency

Using the definition (2.25) of $\hat{\mathbf{V}}^k(\hat{E})$ and the definitions of $\hat{\mathbf{V}}_{RT}^{k-1}(\hat{E})$ and $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$, we have that for $\hat{\mathbf{q}} \in \hat{\mathbf{V}}^k(\hat{E})$,

in 2d: $q_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{P}^k(\hat{y})$, $q_2 \in \mathcal{P}^{k+1}(\hat{y}) \otimes \mathcal{P}^k(\hat{x})$,

in 3d: $q_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z})$, $q_2 \in \mathcal{P}^{k+1}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z})$,

$q_3 \in \mathcal{P}^{k+1}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y})$.

For the degrees of freedom of $\hat{\mathbf{V}}^k$, we consider the following moments:

$$\text{for } k \geq 1, \quad \hat{\mathbf{q}} \mapsto \int_{\hat{e}} \hat{\mathbf{q}} \cdot \hat{\mathbf{n}}_{\hat{e}} \, \hat{p}, \quad \forall \hat{p} \in \mathcal{Q}^k(\hat{e}), \ \forall \hat{e} \in \partial \hat{E}, \tag{2.27}$$

$$\text{for } k \geq 2, \quad \hat{\mathbf{q}} \mapsto \int_{\hat{E}} \hat{\mathbf{q}} \cdot \hat{\mathbf{p}}, \quad \forall \hat{\mathbf{p}} \in \begin{cases} \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{P}^k(\hat{y}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{P}^k(\hat{x}) \end{pmatrix} & \text{in 2d,} \\ \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z}) \\ \mathcal{P}^{k-2}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y}) \end{pmatrix} & \text{in 3d.} \end{cases} \tag{2.28}$$

The number of degrees of freedom given by (2.27) and (2.28) is $2d(k+1)^{d-1}$ and $d(k-1)(k+1)^{d-1}$, respectively. Therefore the total number of DOFs is $d(k+1)^d$, which is same as the $\dim \mathcal{Q}^k(\hat{E}, \mathbb{R}^d)$. We notice that similarly to classical mixed finite elements such as the Raviart–Thomas or Brezzi–Douglas–Marini families of elements, the first set of moments (2.27) stands for facet DOFs, which will be required to be continuous across the facet. The second set of moments (2.28) represents interior DOFs, and no continuity requirements will be imposed on these. These new elements can be viewed as the Raviart–Thomas family with added bubbles, which are curls of specially chosen polynomials.

Theorem 2.2. *Let $\hat{\mathbf{V}}^k(\hat{E})$ be defined as in (2.25). For $\hat{\mathbf{v}} \in \hat{\mathbf{V}}^k(\hat{E})$ suppose that the evaluations of DOFs (2.27) and (2.28) are all zeros. Then $\hat{\mathbf{v}} = 0$.*

Proof. Without loss of generality, we present the proof for $\hat{E} = [-1, 1]^d$. We prove the theorem in 3d, while the 2d result can be obtained in the same manner. From the

definition of shape functions of $\hat{\mathbf{V}}^k(\hat{E})$, $\hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}} \in \mathcal{Q}^k(\hat{e})$ for a face $\hat{e}$ of $\hat{E}$. Therefore, vanishing DOFs (2.27) imply that

$$\hat{\mathbf{v}} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} (1 - \hat{x}^2)\tilde{v}_1(\hat{x}, \hat{y}, \hat{z}) \\ (1 - \hat{y}^2)\tilde{v}_2(\hat{x}, \hat{y}, \hat{z}) \\ (1 - \hat{z}^2)\tilde{v}_3(\hat{x}, \hat{y}, \hat{z}) \end{pmatrix}, \tag{2.29}$$

with

$$\tilde{v}_1 \in \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z}), \quad \tilde{v}_2 \in \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z}), \quad \tilde{v}_3 \in \mathcal{P}^{k-1}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y}).$$

In addition, the vanishing DOFs (2.28) further reduce $\tilde{v}_i$, $i = 1, 2, 3$, to

$$\tilde{v}_1 = L_w^{k-1}(\hat{x})w_1(\hat{y}, \hat{z}), \quad \tilde{v}_2 = L_w^{k-1}(\hat{y})w_2(\hat{x}, \hat{z}), \quad \tilde{v}_3 = L_w^{k-1}(\hat{z})w_3(\hat{x}, \hat{y}), \tag{2.30}$$

where $w_1 \in \mathcal{Q}^k(\hat{y}, \hat{z})$, etc., and $L_w^{k-1}(t)$ is the monic polynomial of degree $k - 1$ on $[-1, 1]$ orthogonal to $\mathcal{P}^{k-2}(t)$ with weight $(1 - t^2)$. Since all monomials in $\hat{\mathbf{V}}^k(\hat{E})$ are of degree $\leq 3k$, $\hat{y}^k \hat{z}^k$ is not contained in $w_1(\hat{y}, \hat{z})$. Similar statements hold with $\hat{z}^k \hat{x}^k$, $\hat{x}^k \hat{y}^k$ and $w_2(\hat{x}, \hat{z})$, $w_3(\hat{x}, \hat{y})$, respectively. Therefore, we can write

$$w_1(\hat{y}, \hat{z}) = \hat{y}^k p_1(\hat{z}) + \hat{z}^k q_1(\hat{y}) + \tilde{w}_1(\hat{y}, \hat{z}),$$

where $p_1 \in \mathcal{P}^{k-1}(\hat{z}), q_1 \in \mathcal{P}^{k-1}(\hat{y}), \tilde{w}_1(\hat{y}, \hat{z}) \in \mathcal{Q}^{k-1}(\hat{y}, \hat{z})$; similar expressions are available for w_2 and w_3. If $p_1 \neq 0$, v_1 has monomials with factor $\hat{x}^{k+1}\hat{y}^k$. From the forms of $\tilde{\boldsymbol{\mathcal{B}}}_i^k(\hat{E})$, $i = 1, 2, 3$, this can be obtained only from a linear combination of elements in $\tilde{\boldsymbol{\mathcal{B}}}_3^k(\hat{E})$ with $c_1 = c_2 = k$. However, a linear combination of elements in $\tilde{\boldsymbol{\mathcal{B}}}_3^k(\hat{E})$ which gives $\hat{x}^{k+1}\hat{y}^k p_1(\hat{z})$ in the first component also has the third component $-(2k+2)\hat{x}^k \hat{y}^k P_1(\hat{z})$ where $P_1(\hat{z})$ is the anti-derivative of $p_1(\hat{z})$ with $P_1(0) = 0$. All terms in v_3 having $\hat{x}^k \hat{y}^k$ as a factor are obtained only from $\tilde{\boldsymbol{\mathcal{B}}}_3^k(\hat{E})$. Furthermore, v_3 does not contain any terms with factor $\hat{x}^k \hat{y}^k$ due to the form of w_3 we discussed, therefore $P_1 = 0$ and $p_1 = 0$ as well. Applying a similar argument we can conclude that $q_1 = 0$, so $w_1 \in \mathcal{Q}^{k-1}(\hat{y}, \hat{z})$. In addition, we can show that $w_2 \in \mathcal{Q}^{k-1}(\hat{x}, \hat{z})$ and $w_3 \in \mathcal{Q}^{k-1}(\hat{x}, \hat{y})$ by similar arguments.

We now claim that $\nabla \cdot \hat{\mathbf{v}} = 0$. First, $\nabla \cdot \hat{\mathbf{v}} \in \mathcal{Q}^{k-1}(\hat{E})$ holds from the definition of the shape functions. Then the Green's identity and the vanishing DOFs assumption give

$$\int_{\hat{E}} \nabla \cdot \hat{\mathbf{v}} q \, d\hat{\mathbf{x}} = \int_{\partial \hat{E}} \hat{\mathbf{v}} \cdot \mathbf{n} \, q \, d\hat{s} - \int_{\hat{E}} \hat{\mathbf{v}} \cdot \nabla q \, d\hat{\mathbf{x}} = 0 \tag{2.31}$$

for any $q \in \mathcal{Q}^{k-1}(\hat{E})$. In particular $q = \nabla \cdot \hat{\mathbf{v}}$ gives $\nabla \cdot \hat{\mathbf{v}} = 0$. From the expression of $\hat{\mathbf{v}}$ in (2.30),

$$0 = \nabla \cdot \hat{\mathbf{v}} = \tilde{L}^k(\hat{x})w_1(\hat{y}, \hat{z}) + \tilde{L}^k(\hat{y})w_2(\hat{x}, \hat{z}) + \tilde{L}^k(\hat{z})w_3(\hat{x}, \hat{y}),$$

where $\tilde{L}^k(t) = \frac{d}{dt}((1 - t^2)L_w^{k-1}(t))$. For $0 \leq i \leq k - 1$, note that

$$\int_{-1}^1 \tilde{L}^k(t)t^i \, dt = -i \int_{-1}^1 (1 - t^2)L_w^{k-1}(t)t^{i-1} \, dt = 0$$

by integration by parts and the definition of L_w^{k-1}. From this observation, we can obtain

$$0 = \int_{\hat{E}} (\nabla \cdot \hat{\mathbf{v}}) \tilde{L}^k(\hat{x}) w_1(\hat{y}, \hat{z}) d\hat{\mathbf{x}} = \int_{\hat{E}} (\tilde{L}^k(\hat{x}) w_1(\hat{y}, \hat{z}))^2 \, d\hat{\mathbf{x}},$$

which implies $w_1 = 0$. We can conclude $w_2 = w_3 = 0$ with similar arguments, therefore $\hat{\mathbf{v}} = 0$. □

2.4.3. *Mixed finite element spaces*

For $k \geq 1$, consider the pair of mixed finite element spaces $\hat{\mathbf{V}}^k(\hat{E}) \times \hat{W}^{k-1}(\hat{E})$, recalling that

$$\hat{\mathbf{V}}^k(\hat{E}) = \hat{\mathbf{V}}_{RT}^{k-1}(\hat{E}) \oplus \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}), \quad \hat{W}^{k-1}(\hat{E}) = \mathcal{Q}^{k-1}(\hat{E}).$$

Note that the construction of $\hat{\mathbf{V}}^k(\hat{E})$ and (2.11) imply that

$$\hat{\nabla} \cdot \hat{\mathbf{V}}^k(\hat{E}) = \hat{W}^{k-1}(\hat{E}) \quad \text{and} \quad \forall \hat{\mathbf{v}} \in \hat{\mathbf{V}}^k(\hat{E}), \quad \forall \hat{e} \subset \partial \hat{E}, \quad \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}} \in \mathcal{Q}^k(\hat{e}). \tag{2.32}$$

Recall also that $\dim \hat{\mathbf{V}}^k(\hat{E}) = \dim \mathcal{Q}^k(\hat{E}, \mathbb{R}^d) = d(k+1)^d$ and that its degrees of freedom are the moments (2.27) and (2.28). We consider an alternative definition of degrees of freedom involving the values of vector components at the Gauss–Lobatto quadrature points; see Fig. 1, where filled arrows indicate the facet degrees of freedom for which continuity across facets is required, and unfilled arrows represent the "interior" degrees of freedom, local to each element. We have omitted some of the degrees of freedom from the backplane of the cube for clarity of visualization. This choice gives certain orthogonalities for the Gauss–Lobatto quadrature rule which we will discuss in details in the forthcoming subsections.
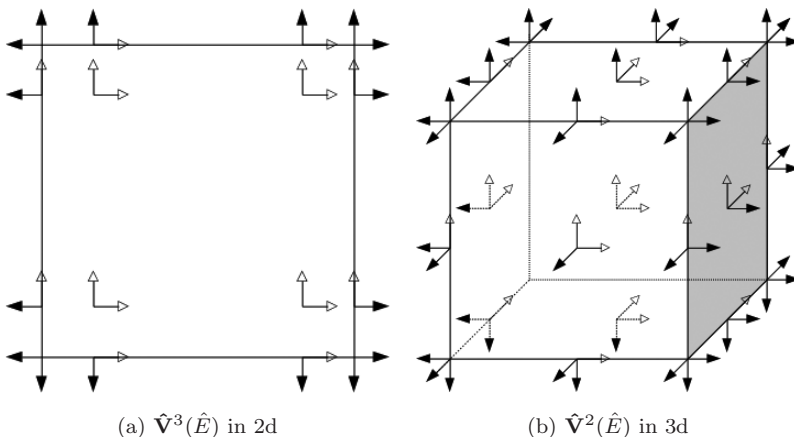

(a) $\hat{\mathbf{V}}^3(\hat{E})$ in 2d   (b) $\hat{\mathbf{V}}^2(\hat{E})$ in 3d

Fig. 1.   Degrees of freedom of the enhanced Raviart–Thomas elements.

The unisolvency of the enhanced Raviart–Thomas spaces shown in the previous section implies the existence of a unique projection operator $\hat{\Pi}_*^k : H^1(\hat{E}, \mathbb{R}^d) \to \hat{\mathbf{V}}^k(\hat{E})$ such that

$$\text{for } k \geq 1, \quad \langle (\hat{\Pi}_*^k \hat{\mathbf{q}} - \hat{\mathbf{q}}) \cdot \mathbf{n}_{\hat{e}}, \, \hat{p} \rangle_{\hat{e}} = 0, \quad \forall \hat{e} \subset \partial \hat{E}, \quad \forall \hat{p}_k \in \mathcal{Q}^k(\hat{e}), \tag{2.33}$$

$$\text{for } k \geq 2, \quad (\hat{\Pi}_*^k \hat{\mathbf{q}} - \hat{\mathbf{q}}, \hat{\mathbf{p}})_{\hat{E}} = 0, \quad \forall \hat{\mathbf{p}} \in \begin{cases} \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{P}^k(\hat{y}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{P}^k(\hat{x}) \end{pmatrix} & \text{in 2d,} \\ \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z}) \\ \mathcal{P}^{k-2}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y}) \end{pmatrix} & \text{in 3d.} \end{cases} \tag{2.34}$$

The Green's identity (2.31) together with (2.33) and (2.34) implies that

$$\left(\hat{\nabla} \cdot (\hat{\Pi}_*^k \hat{\mathbf{q}} - \hat{\mathbf{q}}), \hat{w} \right)_{\hat{E}} = 0, \quad \forall \hat{w} \in \hat{W}^{k-1}(\hat{E}). \tag{2.35}$$

Using (2.15), the above implies that

$$\left(\nabla \cdot (\Pi_*^k \mathbf{q} - \mathbf{q}), w \right)_E = 0, \quad \forall w \in W^{k-1}(E). \tag{2.36}$$

Let $\mathbf{V}_h^k \times W_h^{k-1}$ be the pair of enhanced Raviart–Thomas spaces on $\mathcal{T}_h$ defined as in (2.17) and the projection operator Π_*^k from $\mathbf{V} \cap H^1(\Omega, \mathbb{R}^d)$ onto $\mathbf{V}_h^k$ be defined via the Piola transformation as in (2.18).

Lemma 2.1. *There exists a positive constant β, independent of h, such that*

$$\inf_{w \in W_h^{k-1}} \sup_{\mathbf{q} \in \mathbf{V}_h^k} \frac{(\nabla \cdot \mathbf{q}, w)}{\|w\| \|\mathbf{q}\|_{\text{div}}} \geq \beta. \tag{2.37}$$

Proof. We consider the auxiliary problem

$$\nabla \cdot \boldsymbol{\psi} = w \quad \text{in } \Omega, \quad \boldsymbol{\psi} = \mathbf{g} \quad \text{on } \partial \Omega, \tag{2.38}$$

where $\mathbf{g} \in H^{1/2}(\partial \Omega, \mathbb{R}^d)$ is constructed such that it satisfies $\int_{\partial \Omega} \mathbf{g} \cdot \mathbf{n} = \int_\Omega w$ and $\mathbf{g} \cdot \mathbf{n} = 0$ on Γ_N. More specifically, we choose $\mathbf{g} = (\int_{\partial \Omega} w) \phi \mathbf{n}$, where $\phi \in C^0(\partial \Omega)$ is such that $\int_{\partial \Omega} \phi = 1$ and $\phi = 0$ on Γ_N. Clearly, such construction implies $\|\mathbf{g}\|_{1/2, \partial \Omega} \leq C \|w\|$. It is known[25] that the problem (2.38) has a solution satisfying

$$\|\boldsymbol{\psi}\|_1 \leq C(\|w\| + \|\mathbf{g}\|_{1/2, \partial \Omega}) \leq C \|w\|. \tag{2.39}$$

As the solution $\boldsymbol{\psi}$ is regular enough, $\Pi_*^k \boldsymbol{\psi}$ is well defined. Using (2.36), the choice $\mathbf{q} = \Pi_*^k \boldsymbol{\psi} \in \mathbf{V}_h^k$ yields

$$\left(\nabla \cdot \mathbf{q}, w \right) = \left(\nabla \cdot \Pi_*^k \boldsymbol{\psi}, w \right) = \left(\nabla \cdot \boldsymbol{\psi}, w \right) = \|w\|^2.$$

We complete the proof by exploiting the continuity bound $\|\Pi_*^k \boldsymbol{\psi}\|_{\text{div}} \leq C \|\boldsymbol{\psi}\|_1$, which is stated in (3.22) below. $\square$

We also note that since $\mathbf{V}_{RT,h}^{k-1} \subset \mathbf{V}_h^k$, it follows from the definition of Π_{RT}^k that

$$\nabla \cdot \mathbf{v} = \nabla \cdot \Pi_{RT}^{k-1}\mathbf{v}, \quad \forall \mathbf{v} \in \mathbf{V}_h^k, \tag{2.40}$$

$$\|\Pi_{RT}^{k-1}\mathbf{v}\| \leq C\|\mathbf{v}\|, \quad \forall \mathbf{v} \in \mathbf{V}_h^k. \tag{2.41}$$

2.5. *Quadrature rule*

We next present the quadrature rule for the velocity bilinear form, which is designed to allow for local velocity elimination around finite element nodes. We perform the integration on any element by mapping to the reference element $\hat{E}$. The quadrature rule is defined on $\hat{E}$. We have for $\mathbf{q}, \mathbf{v} \in \mathbf{V}_h^k$,

$$\int_E K^{-1}\mathbf{q} \cdot \mathbf{v} \, d\mathbf{x} = \int_{\hat{E}} \hat{K}^{-1}\frac{1}{J_E}DF_E\hat{\mathbf{q}} \cdot \frac{1}{J_E}DF_E\hat{\mathbf{v}} \, J_E d\hat{\mathbf{x}}$$

$$= \int_{\hat{E}} \frac{1}{J_E}DF_E^T\hat{K}^{-1}DF_E\hat{\mathbf{q}} \cdot \hat{\mathbf{v}} \, d\hat{\mathbf{x}} \equiv \int_{\hat{E}} \mathcal{K}^{-1}\hat{\mathbf{q}} \cdot \hat{\mathbf{v}} \, d\hat{\mathbf{x}},$$

where

$$\mathcal{K} = J_E DF_E^{-1}\hat{K}(DF_E^{-1})^T. \tag{2.42}$$

It is straightforward to show that (2.3) and (2.9) imply that

$$\|\mathcal{K}\|_{0,\infty,\hat{E}} \sim h^{d-2}\|K\|_{0,\infty,E}, \quad \|\mathcal{K}^{-1}\|_{0,\infty,\hat{E}} \sim h^{2-d}\|K^{-1}\|_{0,\infty,E}. \tag{2.43}$$

Let $\Xi_k := \{\xi_k(i)\}_{i=0}^k$ and $\Lambda_k := \{\lambda_k(i)\}_{i=0}^k$ be the points and weights of the Gauss–Lobatto quadrature rule on $[-1,1]$. If k is clear in context, we use $(p,q)_Q$ to denote the evaluation of Gauss–Lobatto quadrature with $k+1$ points for (p,q). We also define

$$\hat{\mathbf{p}}_{\boldsymbol{i}} := (\xi_k(\boldsymbol{i}_1), \ldots, \xi_k(\boldsymbol{i}_d)), \quad w_k(\boldsymbol{i}) := \lambda_k(\boldsymbol{i}_1)\cdots\lambda_k(\boldsymbol{i}_d)$$

$$\text{for } \boldsymbol{i} \in \mathcal{I}_k \equiv \{(\boldsymbol{i}_1, \ldots, \boldsymbol{i}_d), \, \boldsymbol{i}_j \in \{0, \ldots, k\}\}. \tag{2.44}$$

For the method of order k, the quadrature rule is defined on an element E as follows

$$\left(K^{-1}\mathbf{q}, \mathbf{v}\right)_{Q,E} \equiv \left(\mathcal{K}^{-1}\hat{\mathbf{q}}, \hat{\mathbf{v}}\right)_{\hat{Q},\hat{E}} \equiv \sum_{\boldsymbol{i}\in\mathcal{I}_k} w_k(\boldsymbol{i})\mathcal{K}^{-1}(\hat{\mathbf{p}}_{\boldsymbol{i}})\hat{\mathbf{q}}(\hat{\mathbf{p}}_{\boldsymbol{i}}) \cdot \hat{\mathbf{v}}(\hat{\mathbf{p}}_{\boldsymbol{i}}). \tag{2.45}$$

The global quadrature rule can then be defined as

$$\left(K^{-1}\mathbf{q}, \mathbf{v}\right)_Q \equiv \sum_{E\in\mathcal{T}_h} \left(K^{-1}\mathbf{q}, \mathbf{v}\right)_{Q,E}.$$

Note that the method in the lowest order case $k = 1$ is very similar in nature to the one developed in Refs. 27 and 47, although we use different finite element spaces.

We next show that the evaluation at the tensor-product quadrature points is a set of DOFs of $\hat{\mathbf{V}}^k(\hat{E})$, so the bilinear form with the quadrature is not degenerate.

Lemma 2.2. *For $p \in \mathcal{Q}^k(\hat{E})$, if the evaluations of p vanish at all the quadrature nodes of the tensor product Gauss–Lobatto rules on $\hat{E}$, then $p = 0$.*

The above statement is obvious, because the evaluations at the tensor product quadrature nodes become a set of DOFs of $\mathcal{Q}^k(\hat{E})$.

Lemma 2.3. *For* $\hat{\mathbf{q}} \in \hat{\mathbf{V}}^k(\hat{E})$, *if* $\hat{\mathbf{q}}(\hat{\mathbf{p}}_i) = 0$ *for all* $\hat{\mathbf{p}}_i$ *in* (2.44), *then* $\hat{\mathbf{q}} = 0$.

Proof. Without loss of generality, we present the proof for $\hat{E} = [-1, 1]^d$. It suffices to show that the vanishing quadrature evaluation assumption implies that the moments in (2.27) and (2.28) vanish. Since $\hat{\mathbf{q}} \cdot \mathbf{n}_e \in \mathcal{Q}^k(e) \ \forall e \subset \partial\hat{E}$, the vanishing quadrature assumption for nodes on e implies that $\hat{\mathbf{q}} \cdot \mathbf{n}_e = 0$. Therefore, the moments in (2.27) vanish and $\hat{\mathbf{q}}$ is reduced to the form in (2.29), i.e.

$$
\hat{\mathbf{q}} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} (1 - \hat{x}^2)\tilde{q}_1(\hat{x}, \hat{y}, \hat{z}) \\ (1 - \hat{y}^2)\tilde{q}_2(\hat{x}, \hat{y}, \hat{z}) \\ (1 - \hat{z}^2)\tilde{q}_3(\hat{x}, \hat{y}, \hat{z}) \end{pmatrix},
$$

with

$$
\tilde{q}_1 \in \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z}), \quad \tilde{q}_2 \in \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z}), \quad \tilde{q}_3 \in \mathcal{P}^{k-1}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y}).
$$

We want to show that all moments (2.28) of $\hat{\mathbf{q}}$ are zeros. To do it, we first express $\tilde{q}_1$ as

$$
\tilde{q}_1 = \sum_{j=0}^{k-1} L_w^j(\hat{x}) r_j(\hat{y}, \hat{z}), \quad r_j(\hat{y}, \hat{z}) \in \mathcal{Q}^k(\hat{y}, \hat{z}), \tag{2.46}
$$

where L_w^j is the Legendre polynomial of degree j with weight $(1 - \hat{x}^2)$ as before. For fixed $\hat{y}$ and $\hat{z}$, let us consider the Gauss–Lobatto quadrature of $q_1 v$ along $\hat{x}$ with $v \in \mathcal{P}^{k-2}(\hat{x})$. For fixed values of $\hat{y}$ and $\hat{z}$, q_1 is a polynomial of degree $\leq k + 1$, so this quadrature evaluation of $q_1 v$ equals the integration of $q_1 v$ in $\hat{x}$ with the fixed $\hat{y}$ and $\hat{z}$. In particular, if $v = L_w^m(\hat{x})$, $0 \leq m \leq k - 2$, $\hat{y} = \xi_k(i)$, $\hat{z} = \xi_k(j)$, then the vanishing quadrature assumption and the expression of $\tilde{q}_1$ in (2.46) give

$$
0 = \sum_{l=0}^k \lambda_k(l) q_1(\xi_k(l), \xi_k(i), \xi_k(j)) v(\xi_k(l)) = \int_{-1}^1 q_1(\hat{x}, \xi_k(i), \xi_k(j)) v(\hat{x}) d\hat{x}
$$

$$
= \int_{-1}^1 (1 - \hat{x}^2)(L_w^m(\hat{x}))^2 r_m(\xi_k(i), \xi_k(j)).
$$

This implies that $r_m(\hat{y}, \hat{z}) = 0$ for any $\hat{y} = \xi_k(i)$, $\hat{z} = \xi_k(j)$, $0 \leq i, j \leq k$ if $0 \leq m \leq k - 2$, and therefore $r_m = 0$ for $0 \leq m \leq k - 2$ by Lemma 2.2. As a consequence, $q_1 = (1 - \hat{x}^2) L_w^{k-1}(\hat{x}) r_{k-1}(\hat{y}, \hat{z})$ with $r_{k-1} \in \mathcal{Q}^k(\hat{y}, \hat{z})$ and its evaluations at the DOFs given by the first component in (2.28) vanish. We can derive similar results for q_2 and q_3, i.e. $\hat{\mathbf{q}}$ gives only vanishing moments for the DOFs (2.28). We can conclude that $\hat{\mathbf{q}} = 0$ by the same argument as in the previous proof of unisolvency. $\square$

The above result allows us to define a set of DOFs of $\hat{\mathbf{V}}^k(\hat{E})$ as the evaluations of the vectors at the tensor-product quadrature points $\hat{\mathbf{p}}_i$, $i \in \mathcal{I}_k$. Examples were

given in Fig. 1. Recall that for points on $\partial\hat{E}$, some of the vector components are facet degrees of freedom for which continuity across facets is required, while some are "interior" degrees of freedom, local to each element. For convenience of notation, denote the set of points $\hat{\mathbf{p}}_i$ by $\hat{\mathbf{p}}_i$, $i = 1,\ldots,n_k$, $n_k = (k+1)^d$. Any vector $\hat{\mathbf{q}}(\hat{\mathbf{p}}_i)$ at the node $\hat{\mathbf{p}}_i$ is uniquely determined by its d components evaluated at this node. Since we chose the Gauss–Lobatto (or trapezoid, when $k = 1$) quadrature points for the construction of the velocity degrees of freedom, we are guaranteed to have d orthogonal DOFs associated with each node (quadrature point) $\hat{\mathbf{p}}_i$, and they uniquely determine the nodal vector $\hat{\mathbf{q}}(\hat{\mathbf{p}}_i)$. More precisely,

$$\hat{\mathbf{q}}(\hat{\mathbf{p}}_i) = \sum_{j=1}^{d}(\hat{\mathbf{q}}\cdot\hat{\mathbf{n}}_{ij})(\hat{\mathbf{p}}_i)\hat{\mathbf{n}}_{ij}, \tag{2.47}$$

where $\hat{\mathbf{n}}_{ij}$, $j = 1,\ldots,d$, are the outward unit normal vectors to the d hyperplanes of dimension $(d-1)$ that intersect at $\hat{\mathbf{p}}_i$, each one parallel to one of the three mutually orthogonal facets of the reference element. Denote the velocity basis functions associated with $\hat{\mathbf{p}}_i$ by $\hat{\mathbf{v}}_{ij}$, $j = 1,\ldots,d$, i.e.

$$(\hat{\mathbf{v}}_{ij}\cdot\hat{\mathbf{n}}_{ij})(\hat{\mathbf{p}}_i) = 1, \quad (\hat{\mathbf{v}}_{ij}\cdot\hat{\mathbf{n}}_{im})(\hat{\mathbf{p}}_i) = 0, \quad m \neq j, \quad \text{and}$$
$$(\hat{\mathbf{v}}_{ij}\cdot\hat{\mathbf{n}}_{lm})(\hat{\mathbf{p}}_l) = 0, \quad l \neq i, \; m = 1,\ldots,d. \tag{2.48}$$

The quadrature rule (2.45) couples only d basis functions associated with a node. For example, in 3d, for any node $i = 1,\ldots,n_k$,

$$\left(\mathcal{K}^{-1}\hat{\mathbf{v}}_{i1},\hat{\mathbf{v}}_{i1}\right)_{\hat{Q},\hat{E}} = \mathcal{K}_{11}^{-1}(\hat{\mathbf{p}}_i)w_k(i), \quad \left(\mathcal{K}^{-1}\hat{\mathbf{v}}_{i1},\hat{\mathbf{v}}_{i2}\right)_{\hat{Q},\hat{E}} = \mathcal{K}_{21}^{-1}(\hat{\mathbf{p}}_i)w_k(i),$$
$$\left(\mathcal{K}^{-1}\hat{\mathbf{v}}_{i1},\hat{\mathbf{v}}_{i3}\right)_{\hat{Q},\hat{E}} = \mathcal{K}_{31}^{-1}(\hat{\mathbf{p}}_i)w_k(i), \quad \left(\mathcal{K}^{-1}\hat{\mathbf{v}}_{i1},\hat{\mathbf{v}}_{mj}\right)_{\hat{Q},\hat{E}} = 0 \quad \forall mj \neq i1, i2, i3. \tag{2.49}$$

By mapping back (2.45) to the physical element E, we obtain

$$\left(K^{-1}\mathbf{q},\mathbf{v}\right)_{Q,E} = \sum_{i=1}^{n_k}J_E(\hat{\mathbf{p}}_i)w_k(i)K^{-1}(\mathbf{p}_i)\mathbf{q}(\mathbf{p}_i)\cdot\mathbf{v}(\mathbf{p}_i). \tag{2.50}$$

Denote the element quadrature error by

$$\sigma_E\left(K^{-1}\mathbf{q},\mathbf{v}\right) \equiv \left(K^{-1}\mathbf{q},\mathbf{v}\right)_E - \left(K^{-1}\mathbf{q},\mathbf{v}\right)_{Q,E}, \tag{2.51}$$

and define the global quadrature error by $\sigma\left(K^{-1}\mathbf{q},\mathbf{v}\right)\big|_E = \sigma_E\left(K^{-1}\mathbf{q},\mathbf{v}\right)$. Similarly, denote the quadrature error on the reference element by

$$\hat{\sigma}_E\left(\mathcal{K}^{-1}\hat{\mathbf{q}},\hat{\mathbf{v}}\right) \equiv \left(\mathcal{K}^{-1}\hat{\mathbf{q}},\hat{\mathbf{v}}\right)_{\hat{E}} - \left(\mathcal{K}^{-1}\hat{\mathbf{q}},\hat{\mathbf{v}}\right)_{\hat{Q},\hat{E}}. \tag{2.52}$$

The following lemma will be used to bound the quadrature error.

Lemma 2.4. *For any $\hat{\mathbf{q}} \in \hat{\mathbf{V}}^k(\hat{E})$ and for any $k \geq 1$,*

$$\left(\hat{\mathbf{q}} - \hat{\Pi}_{RT}^{k-1}\hat{\mathbf{q}},\hat{\mathbf{v}}\right)_{\hat{Q},\hat{E}} = 0, \quad \text{for all vectors } \hat{\mathbf{v}} \in \mathcal{Q}^{k-1}(\hat{E},\mathbb{R}^d). \tag{2.53}$$

Proof. Without loss of generality, we present the proof for $\hat{E} = [-1, 1]^d$. We show a detailed proof only for the 3d case because the 2d case is similar. Let v_i, $i = 1, 2, 3$ be the ith component of $\hat{\mathbf{q}} - \hat{\Pi}_{RT}^{k-1}\hat{\mathbf{q}}$. Considering the expression v_1 with the basis of Legendre polynomials, the definition of shape functions in $\hat{\mathbf{V}}^k(\hat{E})$ and the constraints from (2.13) yield that v_1 has the form

$$v_1 = L^{k-1}(\hat{x})p_1(\hat{y}, \hat{z}) + L^k(\hat{x})q_1(\hat{y}, \hat{z}) + L^{k+1}(\hat{x})r_1(\hat{y}, \hat{z})$$

$$+ L^k(\hat{y})u_1(\hat{x}, \hat{z}) + L^k(\hat{z})w_1(\hat{x}, \hat{y}), \tag{2.54}$$

where L^i is the standard ith Legendre polynomial as before, $p_1, q_1, r_1 \in \mathcal{Q}^{k-1}(\hat{y}, \hat{z})$,

$$u_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{P}^{k-1}(\hat{z}) + \mathcal{Q}^k(\hat{x}, \hat{z}), \quad w_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{P}^{k-1}(\hat{y}) + \mathcal{Q}^k(\hat{x}, \hat{y}). \tag{2.55}$$

From (2.12), the restrictions of v_1 on $\hat{x} = -1$ and on $\hat{x} = 1$ are orthogonal to $\mathcal{Q}^{k-1}(\hat{y}, \hat{z})$, and it gives two equations

$$p_1 + q_1 + r_1 = 0, \quad p_1 - q_1 + r_1 = 0, \tag{2.56}$$

therefore $q_1 = 0$ and $r_1 = -p_1$. A similar argument can be applied to v_2 and v_3. In summary, we have

$$v_1 = (L^{k-1}(\hat{x}) - L^{k+1}(\hat{x}))p_1(\hat{y}, \hat{z}) + L^k(\hat{y})u_1(\hat{x}, \hat{z}) + L^k(\hat{z})w_1(\hat{x}, \hat{y}), \tag{2.57}$$

$$v_2 = (L^{k-1}(\hat{y}) - L^{k+1}(\hat{y}))p_2(\hat{z}, \hat{x}) + L^k(\hat{z})u_2(\hat{x}, \hat{y}) + L^k(\hat{x})w_2(\hat{y}, \hat{z}), \tag{2.58}$$

$$v_3 = (L^{k-1}(\hat{z}) - L^{k+1}(\hat{z}))p_3(\hat{x}, \hat{y}) + L^k(\hat{x})u_3(\hat{y}, \hat{z}) + L^k(\hat{y})w_3(\hat{y}, \hat{z}), \tag{2.59}$$

where u_2, u_3, w_2, w_3 belong to polynomial spaces similar to the spaces in (2.55) with variable permutation. To prove $(v_1, q)_{\hat{Q}, \hat{E}} = 0$ for $q \in \mathcal{Q}^{k-1}(\hat{E})$, we will show

$$((L^{k-1}(\hat{x}) - L^{k+1}(\hat{x}))p_1(\hat{y}, \hat{z}), q)_{\hat{Q}, \hat{E}} = 0,$$

$$(L^k(\hat{y})u_1(\hat{x}, \hat{z}), q)_{\hat{Q}, \hat{E}} = 0, \tag{2.60}$$

$$(L^k(\hat{z})w_1(\hat{x}, \hat{y}), q)_{\hat{Q}, \hat{E}} = 0.$$

For the first equality, recall that the quadrature points of the Gauss–Lobatto rules are the two endpoints and the zeros of $\frac{d}{dt}L^k(t)$ in $[-1, 1]$. It is clear that $L^{k-1} - L^{k+1}$ vanishes at the two endpoints. In addition, $L^{k-1} - L^{k+1}$ vanishes at the zeros of $\frac{d}{dt}L^k(t)$ in $[-1, 1]$ from the identities

$$(k+1)(L^{k+1} - L^{k-1})(t) = (2k+1)(tL^k(t) - L^{k-1}(t)) = (2k+1)\frac{t^2 - 1}{k}\frac{d}{dt}L^k(t).$$

Therefore, the first equality in (2.60) holds. To prove the second equality in (2.60), let us consider a restriction of the tensor product Gauss–Lobatto rule for fixed quadrature points of $\hat{x}$ and $\hat{z}$. For fixed $\hat{x}$ and $\hat{z}$, the product $L^k(\hat{y})u_1(\hat{x}, \hat{z})q(\hat{x}, \hat{y}, \hat{z})$ is a polynomial in $\hat{y}$ of degree at most $2k-1$, so evaluation of $L^k(\hat{y})u_1(\hat{x}, \hat{z})q(\hat{x}, \hat{y}, \hat{z})$ with the restricted Gauss–Lobatto rule is the same as the integration of the function in $\hat{y}$. However, this integration in $\hat{y}$ is zero because $L^k(\hat{y})$ and $q \in \mathcal{Q}^{k-1}(\hat{x}, \hat{y}, \hat{z})$

are orthogonal. Since $(\cdot, \cdot)_{\hat{Q}, \hat{E}}$ is a sum of these restricted Gauss–Lobatto rules, $(L^k(\hat{y})u_1(\hat{x}, \hat{z}), q)_{\hat{Q}, \hat{E}} = 0$. The third equality in (2.60) follows from the same argument as the second equality. Finally, the same argument can be used for v_2 and v_3, so the assertion is proved. $\square$

2.6. *The kth order MFMFE method*

We first define an appropriate projection to be used in the method for the Dirichlet boundary data g. This is necessary for optimal approximation of the boundary condition term. Moreover, the numerical tests suggest that this is not a purely theoretical artifact, as without the projection we indeed see a deterioration in the rates of convergence. For a facet $\hat{e} \in \partial \hat{E}$, let $\hat{\mathcal{R}}_{\hat{e}}^{k-1}$ be the $L^2(\hat{e})$-orthogonal projection onto $\mathcal{Q}^{k-1}(\hat{e})$, satisfying for any $\hat{\phi} \in L^2(\hat{e})$,

$$\langle \hat{\phi} - \hat{\mathcal{R}}_{\hat{e}}^{k-1} \hat{\phi}, \, \hat{w} \rangle_{\hat{e}} = 0 \quad \forall \hat{w} \in \mathcal{Q}^{k-1}(\hat{e}).$$

Let $\mathcal{R}_h^{k-1} : L^2(\partial \Omega) \to W_h^{k-1}|_{\partial \Omega}$ be such that for any $\phi \in L^2(\partial \Omega)$, $\mathcal{R}_h^{k-1}\phi = \hat{\mathcal{R}}_{\hat{e}}^{k-1}\hat{\phi} \circ F_E^{-1}$ on all $e \in \partial \Omega$. Recall that, c.f. (2.11), if $\hat{\mathbf{v}} \in \hat{\mathbf{V}}_{RT}^{k-1}(\hat{E})$, then $\hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}} \in \mathcal{Q}^{k-1}(\hat{e})$ for all $\hat{e} \subset \partial \hat{E}$. Then, using (2.12) and (2.15), we have that

$$\forall \phi \in L^2(\partial \Omega), \quad \langle \phi - \mathcal{R}_h^{k-1}\phi, \, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial \Omega} = 0 \quad \forall \mathbf{v} \in \hat{\mathbf{V}}_{RT}^{k-1}(\hat{E}) \tag{2.61}$$

and

$$\forall \mathbf{v} \in H^1(\Omega, \mathbb{R}^d), \quad \langle (\mathbf{v} - \Pi_{RT}^{k-1}\mathbf{v}) \cdot \mathbf{n}, \, \mathcal{R}_h^{k-1}\phi \rangle_{\partial \Omega} = 0 \quad \phi \in L^2(\partial \Omega). \tag{2.62}$$

The method is defined as follows: find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h^k \times W_h^{k-1}$, where $k \geq 1$, such that

$$\left(K^{-1}\mathbf{u}_h, \mathbf{v}\right)_Q - \left(p_h, \nabla \cdot \mathbf{v}\right) = -\langle \mathcal{R}_h^{k-1}g, \mathbf{v} \cdot \mathbf{n} \rangle_{\Gamma_D}, \quad \mathbf{v} \in \mathbf{V}_h^k, \tag{2.63}$$

$$\left(\nabla \cdot \mathbf{u}_h, w\right) = \left(f, w\right), \quad w \in W_h^{k-1}. \tag{2.64}$$

Following the terminology from Refs. 27 and 47 we call the method (2.63)–(2.64) a kth order MFMFE method, due to its relation to the MPFA scheme.

In order to prove that the method stated above has a unique solution, we first present several useful results.

Lemma 2.5. *If $E \in \mathcal{T}_h$ and $\mathbf{q} \in L^2(E, \mathbb{R}^d)$, then*

$$\|\mathbf{q}\|_E \sim h^{\frac{2-d}{2}} \|\hat{\mathbf{q}}\|_{\hat{E}}. \tag{2.65}$$

Proof. The statement of the lemma follows from (2.14):

$$\int_E \mathbf{q} \cdot \mathbf{q} \, d\mathbf{x} = \int_{\hat{E}} \frac{1}{J_E} DF_E \hat{\mathbf{q}} \cdot \frac{1}{J_E} DF_E \hat{\mathbf{q}} \, J_E d\hat{\mathbf{x}},$$

$$\int_{\hat{E}} \hat{\mathbf{q}} \cdot \hat{\mathbf{q}} \, d\hat{\mathbf{x}} = \int_E \frac{1}{J_{F_E^{-1}}} DF_E^{-1}\mathbf{q} \cdot \frac{1}{J_{F_E^{-1}}} DF_E^{-1}\mathbf{q} \, J_{F_E^{-1}} d\mathbf{x},$$

and bounds (2.9). $\square$

Lemma 2.6. *The bilinear form* $\left(K^{-1}\mathbf{q}, \mathbf{v}\right)_Q$ *is an inner product on* $\mathbf{V}_h^k$ *and* $\left(K^{-1}\mathbf{q}, \mathbf{q}\right)_Q^{1/2}$ *is a norm in* $\mathbf{V}_h^k$ *equivalent to* $\|\cdot\|$.

Proof. Let $\mathbf{q} \in \mathbf{V}_h^k$ be given on an element E as $\mathbf{q} = \sum_{i=1}^{n_k} \sum_{j=1}^{d} q_{ij}\mathbf{v}_{ij}$. Using (2.3), (2.9), (2.50), and the basis property (2.48), we obtain

$$\left(K^{-1}\mathbf{q}, \mathbf{q}\right)_{Q,E} = \sum_{i=1}^{n_k} J_E(\hat{\mathbf{p}}_i)w_k(i)K^{-1}(\mathbf{p}_i)\mathbf{q}(\mathbf{p}_i) \cdot \mathbf{q}(\mathbf{p}_i) \geq Ch^d \sum_{i=1}^{n_k} \sum_{j=1}^{d} q_{ij}^2.$$

On the other hand,

$$\|\mathbf{q}\|_E^2 = \left(\sum_{i=1}^{n_k} \sum_{j=1}^{d} q_{ij}\mathbf{v}_{ij}, \sum_{k=1}^{n_k} \sum_{l=1}^{d} q_{kl}\mathbf{v}_{kl}\right) \leq Ch^d \sum_{i=1}^{n_k} \sum_{j=1}^{d} q_{ij}^2.$$

Hence,

$$\left(K^{-1}\mathbf{q}, \mathbf{q}\right)_Q \geq C\|\mathbf{q}\|^2, \tag{2.66}$$

and due to the linearity and symmetry, we conclude that $\left(K^{-1}\mathbf{q}, \mathbf{v}\right)_Q$ is an inner product and $\left(K^{-1}\mathbf{q}, \mathbf{q}\right)_Q^{1/2}$ is a norm in $\mathbf{V}_h^k$. Using (2.3), (2.43) (2.45), (2.65), and the equivalence of norms on $\hat{E}$, we obtain

$$\left(K^{-1}\mathbf{q}, \mathbf{q}\right)_{Q,E} = \sum_{\mathbf{i}\in\mathcal{I}_k} w_k(\mathbf{i})\mathcal{K}^{-1}(\hat{\mathbf{p}}_{\mathbf{i}})\hat{\mathbf{q}}(\hat{\mathbf{p}}_{\mathbf{i}}) \cdot \hat{\mathbf{q}}(\hat{\mathbf{p}}_{\mathbf{i}}) \leq Ch^{2-d}\|\hat{\mathbf{q}}\|_{\hat{E}}^2 \leq C\|\mathbf{q}\|_E^2. \tag{2.67}$$

Combining (2.66) and (2.67) results in the equivalence of norms

$$c_0\|\mathbf{q}\| \leq \left(K^{-1}\mathbf{q}, \mathbf{q}\right)_Q^{1/2} \leq c_1\|\mathbf{q}\|. \tag{2.68}$$

$\square$

We now proceed with the solvability of the method (2.63)–(2.64).

Theorem 2.3. *The kth order MFMFE method* (2.63)–(2.64) *has a unique solution for any $k \geq 1$.*

Proof. Since (2.63)–(2.64) is a square system, it is enough to prove uniqueness of the solution. Letting $f = 0$, $g = 0$ and choosing $\mathbf{v} = \mathbf{u}_h$ and $w = p_h$, one immediately obtains $\left(K^{-1}\mathbf{u}_h, \mathbf{u}_h\right)_Q = 0$, which yields $\mathbf{u}_h = 0$ due to (2.68). Next, we use the inf-sup condition (2.37) to obtain

$$\|p_h\| \leq C \sup_{\mathbf{q}\in\mathbf{V}_h^k} \frac{(\nabla \cdot \mathbf{q}, p_h)}{\|\mathbf{q}\|_{\text{div}}} = \sup_{\mathbf{q}\in\mathbf{V}_h^k} \frac{\left(K^{-1}\mathbf{u}_h, \mathbf{q}\right)_Q}{\|\mathbf{q}\|_{\text{div}}} = 0$$

and thus $p_h = 0$, which concludes the proof of the theorem. $\square$

2.7. *Reduction to a pressure system and its stencil*

In this section, we describe how the MFMFE method reduces to a system for the pressures by local velocity elimination. Recall that the DOFs of $\hat{\mathbf{V}}^k(\hat{E})$ are chosen as the d vector components at the tensor-product Gauss–Lobatto quadrature points, see Fig. 1. As a result, in the velocity mass matrix obtained from the bilinear form $(K^{-1}\mathbf{u}_h, \mathbf{v})$, the d DOFs associated with a quadrature point in an element E are completely decoupled from other DOFs in E, see (2.49). Due to the continuity of normal components across facets, there are couplings with DOFs from neighboring elements. We distinguish three types of velocity couplings. The first involves localization of degrees of freedom around each vertex in the grid. Only this type occurs in the lowest order case $k = 1$, similar to the previously developed lowest order MFMFE method.[27,47] The number of DOFs that are coupled around a vertex equals the number of facets n_v that share the vertex. For example, on logically rectangular grids, $n_v = 12$ (faces) in 3d and $n_v = 4$ (edges) in 2d. The second type of coupling is around nodes located on facets, but not at vertices. In 2d, these are edge DOFs. The number of coupled DOFs is three — one normal to the edge, which is continuous across the edge, and two tangential to the edge, one from each of the two neighboring elements. In 3d, there are two cases to consider for this type of coupling. One case is for nodes located on faces, but not on edges. In this case the number of coupled DOFs is five — one normal to the face, which is continuous across the face, and four tangential to the face, two from each of the two neighboring elements. The second case in 3d is for nodes located on edges, but not at vertices. Let n_e be the number of elements that share the edge, which also equals the number of faces that share the edge. In this case the number of coupled DOFs is $2n_e$. These include n_e DOFs normal to the n_e faces, which are continuous across the faces, and n_e DOFs tangential to the edge, one per each of the n_e neighboring elements. For example, on logically rectangular grids, $n_e = 4$, resulting in eight coupled DOFs. Finally, the third type of coupling involves nodes interior to the elements, in which case only the d DOFs associated with the node are coupled.

Due to the localization of DOF interactions described above, the velocity mass matrix obtained from the bilinear form $(K^{-1}\mathbf{u}_h, \mathbf{v})$, is block-diagonal with blocks associated with the Gauss–Lobatto quadrature points. In particular, in 2d, there are $n_v \times n_v$ blocks at vertices (n_v is the number of neighboring edges), 3×3 blocks at edge points, and 2×2 blocks at interior points. In 3d, there are $n_v \times n_v$ blocks at vertices (n_v is the number of neighboring faces), $2n_e \times 2n_e$ blocks at edge points (n_e is the number of neighboring elements), 5×5 blocks at face points, and 3×3 blocks at interior points.

Proposition 2.1. *The local matrices described above are symmetric and positive definite.*

Proof. For any quadrature point, the local matrix is obtained by taking $\mathbf{v} = \mathbf{v}_1, \ldots, \mathbf{v}_m$ in (2.63), where $\mathbf{v}_i$ are the velocity basis functions associated with that

point. We have

$$\left(K^{-1}\mathbf{u}_h, \mathbf{v}_i\right)_Q = \sum_{j=1}^{m} u_j \left(K^{-1}\mathbf{v}_j, \mathbf{v}_i\right) \equiv \sum_{j=1}^{m} a_{ij} u_j, \quad i = 1, \ldots, m.$$

Using Lemma 2.6 we conclude that the matrix $M = \{a_{ij}\}$ is symmetric and positive definite. □

The block-diagonal structure of the velocity mass matrix allows for local velocity elimination. In particular, solving the local linear systems resulting from (2.63) allows us to express the associated velocities in terms of the pressures from the neighboring elements and boundary data. This implies that the method reduces the saddle-point problem to an element-based pressure system.

Lemma 2.7. *The pressure system resulting from* (2.63)–(2.64) *using the procedure described above is symmetric and positive definite.*

Proof. The proof follows from the argument presented in Proposition 2.8 in Ref. 47. We present it here for the sake of completeness. Denoting the bases of $\mathbf{V}_h^k$ and W_h^{k-1} by $\{\mathbf{v}_i\}$ and $\{w_i\}$, respectively, we obtain the saddle-point type algebraic system arising from (2.63)–(2.64),

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} G \\ F \end{pmatrix}, \tag{2.69}$$

where $A_{ij} = \left(K^{-1}\mathbf{v}_i, \mathbf{v}_j\right)_Q$ and $B_{ij}^T = -(\nabla \cdot \mathbf{v}_i, w_j)$. The matrix A obtained by the above procedure is symmetric and positive definite, as it is block diagonal with SPD blocks associated with quadrature nodes shown in Proposition 2.1. The elimination of U leads to a system for P with a symmetric and positive semidefinite matrix $BA^{-1}B^T$. It follows immediately from the proof of Theorem 2.3 that $B^T P = 0$ if and only if $P = 0$. Therefore, $BA^{-1}B^T$ is positive definite. □

Remark 2.1. We note that while $\mathbf{V}_h^k$ has more DOFs than $\mathbf{V}_{RT,h}^{k-1}$ with comparable accuracy, cf. Sec. 3, the above reduction technique allows for local elimination of all velocity DOFs, resulting in a symmetric and positive definite system only for the pressure DOFs in W_h^{k-1}. This is computationally more efficient than solving a saddle point problem for the classical Raviart–Thomas MFE method in $\mathbf{V}_{RT,h}^{k-1} \times W_h^{k-1}$.

Remark 2.2. It was pointed out by an anonymous reviewer that it can be shown that the matrix B has a tensor product structure if it is formed using the tensor product Gauss quadrature rule. This property can be exploited for faster and low storage matrix-free assembly and application of the matrix $BA^{-1}B^T$, resulting in further gain in efficiency. We thank the reviewer for noting this important property.

3. Velocity Error Analysis

Although the proposed schemes can be defined and are well posed on general quadrilateral or hexahedra, for the convergence analysis we need to impose a restriction on the element geometry. This is due to the reduced approximation properties of the MFE spaces on arbitrary shaped quadrilaterals or hexahedra that our new family of elements inherits as well. The necessity of said restriction is confirmed by the numerical computations. We recall that, since the mapping F_E is trilinear in 3d, the faces of an element E may be non-planar. We will refer to the faces as generalized quadrilaterals. We recall the notation of $\mathbf{r}_i$, $i = 1, \ldots, 2^d$, and edges $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ from Sec. 2.2.

Definition 3.1. A (generalized) quadrilateral with vertices $\mathbf{r}_i$, $i = 1, \ldots, 4$, is called an h^2-parallelogram if

$$|\mathbf{r}_{34} - \mathbf{r}_{21}|_{\mathbb{R}^d} \leq Ch^2.$$

The name follows the terminology from Refs. 24 and 27. Note that elements of this type in 2d can be obtained by uniform refinements of a general quadrilateral grid. It follows from (2.6) that $\frac{\partial^2 F_E}{\partial \hat{x} \partial \hat{y}}$ is $\mathcal{O}(h^2)$ for h^2-parallelograms.

Definition 3.2. A hexahedral element is called an h^2-parallelepiped if all of its faces are h^2-parallelograms.

Definition 3.3. An h^2-parallelepiped with vertices $\mathbf{r}_i$, $i = 1, \ldots, 8$, is called regular if

$$|(\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78})|_{\mathbb{R}^3} \leq Ch^3.$$

It is clear from (2.7) that for h^2-parallelepipeds, $\frac{\partial^2 F_E}{\partial \hat{x} \partial \hat{y}}$, $\frac{\partial^2 F_E}{\partial \hat{y} \partial \hat{z}}$ and $\frac{\partial^2 F_E}{\partial \hat{x} \partial \hat{z}}$ are $\mathcal{O}(h^2)$. Moreover, in case of regular h^2-parallelepipeds, $\frac{\partial^3 F_E}{\partial \hat{x} \partial \hat{y} \partial \hat{z}}$ is $\mathcal{O}(h^3)$.

We next present some bounds on the derivatives of the mapping F_E.

Lemma 3.1. *Let $j \geq 0$. Then the bounds*

$$|J_E|_{j,\infty,\hat{E}} \leq Ch^{j+d}, \quad j \leq \alpha, \quad \text{where } \alpha = 1 \text{ in 2d},$$

$$\alpha = 4 \text{ in 3d}, \quad |J_E|_{j,\infty,\hat{E}} = 0, \ j > \alpha, \tag{3.1}$$

and

$$|DF_E|_{j,\infty,\hat{E}} \leq \begin{cases} Ch^{j+1}, & j < d, \\ 0, & j \geq d, \end{cases} \quad \left|\frac{1}{J_E} DF_E\right|_{j,\infty,\hat{E}} \leq Ch^{j-d+1},$$

$$|J_E DF_E^{-1}|_{j,\infty,\hat{E}} \leq \begin{cases} Ch^{j+d-1}, & j \leq d \\ 0, & j > d \end{cases}, \tag{3.2}$$

hold if E is an h^2-parallelogram or a regular h^2-parallelepiped. Moreover, the estimates (3.2) hold for $j = 0$ if E is a general quadrilateral or hexahedron and for $j = 0, 1$ if E is an h^2-parallelepiped.

Proof. We begin with the proof of (3.1). In 2d, (2.6) gives

$$DF_E = [\mathbf{r}_{21}, \mathbf{r}_{41}] + [(\mathbf{r}_{34} - \mathbf{r}_{21})\hat{y}, (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x}],$$

from which it can be shown easily that J_E is a linear function satisfying (3.1). In 3d, (2.7) gives

$$
\begin{aligned}
DF_E = [&\mathbf{r}_{21} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{y} + (\mathbf{r}_{65} - \mathbf{r}_{21})\hat{z} + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{y}\hat{z}; \\
&\mathbf{r}_{41} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x} + (\mathbf{r}_{85} - \mathbf{r}_{41})\hat{z} + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{x}\hat{z}; \\
&\mathbf{r}_{51} + (\mathbf{r}_{65} - \mathbf{r}_{21})\hat{x} + (\mathbf{r}_{85} - \mathbf{r}_{41})\hat{y} + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{x}\hat{y}].
\end{aligned}
$$

$$(3.3)$$

It can be verified that J_E is a polynomial of three variables of total power at most 4 with

$$(J_E)_{\hat{x}\hat{x}\hat{x}} = (J_E)_{\hat{y}\hat{y}\hat{y}} = (J_E)_{\hat{z}\hat{z}\hat{z}} = 0, \tag{3.4}$$

and it can be written as $J_E = \sum_{0 \le r_1 + r_2 + r_3 \le 4} \alpha_{r_1 r_2 r_3} \hat{x}^{r_1} \hat{y}^{r_2} \hat{z}^{r_3}$, where

$$|\alpha_{r_1 r_2 r_3}| \le Ch^{r_1 + r_2 + r_3 + 3}, \tag{3.5}$$

from which (3.1) follows immediately.

We proceed with the proof of (3.2). If E is a general quadrilateral or hexahedron, the bounds with $j = 0$ are stated in (2.9). The estimates in 2d and for $j = 1, 2$ in 3d were shown in Refs. 24, 27 and 47. We now focus on the case when E is a regular h^2-parallelepiped and $j > 2$. Since F_E is bilinear, $|DF_E|_{k,\infty,\hat{E}} = 0$, $\forall k > 2$, and (3.3) gives

$$|DF_E|_{k,\infty,\hat{E}} \le Ch^{k+1}, \quad k = 0, 1, 2. \tag{3.6}$$

Therefore, it follows from the product rule that for any $j > 2$,

$$
\begin{aligned}
\left| \frac{1}{J_E} DF_E \right|_{j,\infty,\hat{E}} \le C \Bigg(&\left| \frac{1}{J_E} \right|_{j,\infty,\hat{E}} |DF_E|_{0,\infty,\hat{E}} + \left| \frac{1}{J_E} \right|_{j-1,\infty,\hat{E}} |DF_E|_{1,\infty,\hat{E}} \\
&+ \left| \frac{1}{J_E} \right|_{j-2,\infty,\hat{E}} |DF_E|_{2,\infty,\hat{E}} \Bigg).
\end{aligned}
\tag{3.7}
$$

We further compute the derivatives of $\dfrac{1}{J_E}$:

$$\left(\frac{1}{J_E} \right)_{\hat{x}} = -\frac{1}{J_E^2} (J_E)_{\hat{x}}, \quad \left(\frac{1}{J_E} \right)_{\hat{x}\hat{x}\hat{x}} = -\frac{6}{J_E^4} (J_E)_{\hat{x}}^3 + \frac{6}{J_E^3} (J_E)_{\hat{x}} (J_E)_{\hat{x}\hat{x}},$$

$$\left(\frac{1}{J_E} \right)_{\hat{x}\hat{x}} = \frac{2}{J_E^3} (J_E)_{\hat{x}}^2 - \frac{1}{J_E^2} (J_E)_{\hat{x}\hat{x}}, \quad \left(\frac{1}{J_E} \right)_{\hat{x}\hat{y}} = \frac{2}{J_E^3} (J_E)_{\hat{x}} (J_E)_{\hat{y}} - \frac{1}{J_E^2} (J_E)_{\hat{x}\hat{y}},$$

$$\left(\frac{1}{J_E} \right)_{\hat{x}\hat{x}\hat{y}} = -\frac{6}{J_E^4} (J_E)_{\hat{x}}^2 (J_E)_{\hat{y}} + \frac{4}{J_E^3} (J_E)_{\hat{x}} (J_E)_{\hat{x}\hat{y}} + \frac{2}{J_E^3} (J_E)_{\hat{y}} (J_E)_{\hat{x}\hat{x}} - \frac{1}{J_E^2} (J_E)_{\hat{x}\hat{x}\hat{y}}$$

$$\left(\frac{1}{J_E}\right)_{\hat{x}\hat{y}\hat{z}} = -\frac{6}{J_E^4}(J_E)_{\hat{x}}(J_E)_{\hat{y}}(J_E)_{\hat{z}} + \frac{2}{J_E^3}(J_E)_{\hat{x}\hat{z}}(J_E)_{\hat{y}}$$

$$+ \frac{2}{J_E^3}(J_E)_{\hat{x}}(J_E)_{\hat{y}\hat{z}} + \frac{2}{J_E^3}(J_E)_{\hat{z}}(J_E)_{\hat{x}\hat{y}} - \frac{1}{J_E^2}(J_E)_{\hat{x}\hat{y}\hat{z}},$$

$$\left(\frac{1}{J_E}\right)_{\hat{x}\hat{x}\hat{y}\hat{z}} = \frac{24}{J_E^5}(J_E)_{\hat{x}}^2(J_E)_{\hat{y}}(J_E)_{\hat{z}} - \frac{12}{J_E^4}(J_E)_{\hat{x}}(J_E)_{\hat{y}}(J_E)_{\hat{x}\hat{z}} - \frac{6}{J_E^4}(J_E)_{\hat{x}}^2(J_E)_{\hat{y}\hat{z}}$$

$$- \frac{12}{J_E^4}(J_E)_{\hat{x}}(J_E)_{\hat{z}}(J_E)_{\hat{x}\hat{y}} + \frac{4}{J_E^3}(J_E)_{\hat{x}\hat{z}}(J_E)_{\hat{x}\hat{y}} + \frac{4}{J_E^3}(J_E)_{\hat{x}}(J_E)_{\hat{x}\hat{y}\hat{z}}$$

$$- \frac{6}{J_E^4}(J_E)_{\hat{z}}(J_E)_{\hat{y}}(J_E)_{\hat{x}\hat{x}} + \frac{2}{J_E^3}(J_E)_{\hat{x}\hat{x}}(J_E)_{\hat{y}\hat{z}} + \frac{2}{J_E^3}(J_E)_{\hat{y}}(J_E)_{\hat{x}\hat{x}\hat{z}}$$

$$+ \frac{2}{J_E^3}(J_E)_{\hat{z}}(J_E)_{\hat{x}\hat{x}\hat{y}} - \frac{1}{J_E^2}(J_E)_{\hat{x}\hat{x}\hat{y}\hat{z}}.$$

We note that due to (3.4) the higher order partial derivatives will consist of the same partials that appear above, while the power of J_E in the denominator will continue to grow. Therefore, it follows from (3.5) that $|\frac{1}{J_E}|_{k,\infty,\hat{E}} \le Ch^{k-3}$, which, combined with (3.6) and (3.7), implies that

$$\left|\frac{1}{J_E}DF_E\right|_{j,\infty,\hat{E}} \le C\left(h^{j-3}h + h^{j-4}h^2 + h^{j-5}h^3\right) \le Ch^{j-2}.$$

To show the last inequality in (3.2), we note that using the cofactor formula for inverse of a matrix, one can verify that $J_E DF_E^{-1}$ is of total degree 3, which implies that for every $k > 3$, $|J_E DF_E^{-1}|_{k,\infty,\hat{E}} = 0$. We also compute

$$((J_E DF_E^{-1})_{11})_{\hat{x}\hat{x}\hat{y}}$$

$$= 2[(y_1 - y_2) + (y_3 - y_4)][(z_5 - z_6) + (z_7 - z_8) + (z_2 - z_1) + (z_4 - z_3)]$$

$$+ 2[(z_1 - z_2) + (z_3 - z_4)][(y_6 - y_5) + (y_8 - y_7) + (y_1 - y_2) + (y_3 - y_4)],$$

with similar expressions for the rest of partial derivatives. Therefore $|J_E D F_E^{-1}|_{3,\infty,\hat{E}} \le Ch^5$.   □

The above bounds allow us to control the norms of the velocity and permeability on the reference element.

Lemma 3.2. *For all* $\mathbf{q} \in H^j(E)$, *there exists a constant C independent of h such that the bound*

$$|\hat{\mathbf{q}}|_{j,\hat{E}} \le Ch^{j+\frac{d-2}{2}}\|\mathbf{q}\|_{j,E} \tag{3.8}$$

holds for every $j \ge 0$ if E is an h^2-parallelogram or regular h^2-parallelepiped, for $j = 0,1$ if E is an h^2-parallelepiped and for $j = 0$ if E is a general quadrilateral or hexahedron.

Proof. The result in 2d was shown in Refs. 24, 47, while the cases $j = 0, 1, 2$ in 3d were proven in Ref. 27. It then suffices to prove the case $j \geq 3$ for regular h^2-parallelepipeds. Let

$$\tilde{\mathbf{q}} = \mathbf{q} \circ F_E(\hat{\mathbf{x}}), \quad \hat{\mathbf{q}} = J_E DF_E^{-1}\tilde{\mathbf{q}}.$$

As it was shown in the previous lemma $|J_E DF_E^{-1}|_{4,\infty,\hat{E}} = 0$, hence (3.2) implies that for $r \geq 3$,

$$|\hat{\mathbf{q}}|_{r,\hat{E}} \leq C\big(h^2|\tilde{\mathbf{q}}|_{r,\hat{E}} + h^3|\tilde{\mathbf{q}}|_{r-1,\hat{E}} + h^4|\tilde{\mathbf{q}}|_{r-2,\hat{E}} + h^5|\tilde{\mathbf{q}}|_{r-3,\hat{E}}\big). \tag{3.9}$$

By change of variables and the chain rule, we have that $|\tilde{\mathbf{q}}|_{j,\hat{E}} \leq Ch^{j-3/2}\|\mathbf{q}\|_{j,E}$, which, combined with (3.9), completes the proof. $\qquad\square$

Lemma 3.3. *There exists a constant C independent of h such that the bound*

$$|\mathcal{K}^{-1}|_{j,\infty,\hat{E}} \leq Ch^{j-d+2}\|K^{-1}\|_{j,\infty,E}. \tag{3.10}$$

holds with $j \geq 0$ on h^2-parallelograms and regular h^2-parallelepipeds, with $j = 0, 1$ on h^2-parallelepipeds and with $j = 0$ on general quadrilaterals and hexahedra.

Proof. The above result with $j = 0$ was already stated in (2.43). Moreover, for $j = 1, 2$ (3.10) was shown in Refs. 27 and 47, so we focus on the case $j \geq 3$ for h^2-parallelograms and regular h^2-parallelepipeds. By the use of a change of variables, the chain rule, and (3.2), it is easy to see that

$$|\hat{K}^{-1}|_{j,\infty,\hat{E}} \leq Ch^j|K^{-1}|_{j,\infty,E}. \tag{3.11}$$

Using (3.2) and the definition of $\mathcal{K}^{-1}$ given in (2.42), we have

$$|\mathcal{K}^{-1}|_{j,\infty,\hat{E}} \leq C \sum_{\substack{0 \leq \alpha,\beta,\gamma \leq j \\ \alpha+\beta+\gamma=j}} \left|\frac{1}{J_E}DF_E\right|_{\alpha,\infty,\hat{E}} |\hat{K}^{-1}|_{\beta,\infty,\hat{E}}|DF_E|_{\gamma,\infty,\hat{E}}$$

$$\leq C \sum_{\substack{0 \leq \alpha,\beta,\gamma \leq j \\ \alpha+\beta+\gamma=j}} h^{\alpha-d+1}h^\beta h^{\gamma+1}\|K^{-1}\|_{j,\infty,E} \leq Ch^{j-d+2}\|K^{-1}\|_{j,\infty,E},$$

where we also used (3.11) for the second inequality. $\qquad\square$

Lemma 3.4. *There exists a constant C independent of h such that on h^2-parallelograms and regular h^2-parallelepipeds*

$$\|\mathbf{q} - \Pi_*^k\mathbf{q}\| + \|\mathbf{q} - \Pi_{RT}^{k-1}\mathbf{q}\| \leq Ch^j\|\mathbf{q}\|_j, \tag{3.12}$$

$$\|\mathbf{q} - \Pi_*^k\mathbf{q}\| \leq Ch^{j+1}\|\mathbf{q}\|_{j+1}, \tag{3.13}$$

$$\|\nabla \cdot (\mathbf{q} - \Pi_*^k\mathbf{q})\| + \|\nabla \cdot (\mathbf{q} - \Pi_{RT}^{k-1}\mathbf{q})\| \leq Ch^j\|\nabla \cdot \mathbf{q}\|_j, \tag{3.14}$$

for $1 \leq j \leq k$. Moreover, (3.12) and (3.14) also hold on h^2-parallelepipeds with $j = 1$.

Proof. We present the proof for Π_*^k only, as the argument for Π_{RT}^{k-1} is similar. Using (2.65) and (3.8), we have

$$\|\mathbf{q} - \Pi_*^k \mathbf{q}\|_E \leq Ch^{\frac{d-2}{2}} \|\hat{\mathbf{q}} - \hat{\Pi}_*^k \hat{\mathbf{q}}\|_{\hat{E}} \leq Ch^{\frac{d-2}{2}} |\hat{\mathbf{q}}|_{j+1,\hat{E}} \leq Ch^{j+1} \|\mathbf{q}\|_{j,E},$$

where $1 \leq j \leq k$. For the second inequality in the above, we used the fact that $\hat{\Pi}_*^k$ preserves all polynomials of degree up to k, i.e. $\mathcal{P}^k(\hat{E}) \subset \hat{\mathbf{V}}^k(\hat{E})$, and applied the Bramble–Hilbert lemma.[19] Summing over the elements completes the proof of the first two statements of the lemma.

For the last inequality, it follows from (2.14) that

$$\int_E \left(\nabla \cdot (\mathbf{q} - \Pi_*^k \mathbf{q})\right)^2 d\mathbf{x} = \int_{\hat{E}} \frac{1}{J_E^2} \left(\hat{\nabla} \cdot (\hat{\mathbf{q}} - \hat{\Pi}_*^k \hat{\mathbf{q}})\right)^2 J_E \, d\hat{\mathbf{x}} \leq Ch^{-d} |\hat{\nabla} \cdot \hat{\mathbf{q}}|_{j,\hat{E}}^2,$$

$$(3.15)$$

where we have used (2.9), (2.35), and the Bramble–Hilbert lemma in the inequality. We also have

$$|\hat{\nabla} \cdot \hat{\mathbf{q}}|_{j,\hat{E}} = |J_E \widehat{\nabla \cdot \mathbf{q}}|_{j,\hat{E}} \leq C \sum_{i=0}^j |J_E|_{i,\infty,\hat{E}} |\widehat{\nabla \cdot \mathbf{q}}|_{j-i,\hat{E}}$$

$$\leq C \sum_{i=0}^j h^{i+d} h^{j-i-\frac{d}{2}} |\nabla \cdot \mathbf{q}|_{j-i,E} \leq Ch^{j+\frac{d}{2}} \|\nabla \cdot \mathbf{q}\|_{j,E}, \qquad (3.16)$$

where we used (3.1) and change of variables back to E in the second inequality. A combination of (3.15) and (3.16), and a summation over all elements completes the proof of (3.14). □

Let $\hat{\mathcal{Q}}^{k-1}$ be the $L^2(\hat{E})$-orthogonal projection onto $\hat{W}^{k-1}(\hat{E})$, satisfying for any $\hat{\phi} \in L^2(\hat{E})$,

$$\left(\hat{\phi} - \hat{\mathcal{Q}}^{k-1}\hat{\phi}, \hat{w}\right)_{\hat{E}} = 0 \quad \forall \hat{w} \in \hat{W}^{k-1}(\hat{E}).$$

Let $\mathcal{Q}_h^{k-1} : L^2(\Omega) \to W_h^{k-1}$ be the projection operator, satisfying for any $\phi \in L^2(\Omega)$,

$$\mathcal{Q}_h^{k-1}\phi = \hat{\mathcal{Q}}^{k-1}\hat{\phi} \circ F_E^{-1} \quad \text{on all } E.$$

It follows from (2.32) that

$$\left(\phi - \mathcal{Q}_h^{k-1}\phi, \nabla \cdot \mathbf{v}\right) = 0 \quad \forall \mathbf{v} \in \mathbf{V}_h^k. \qquad (3.17)$$

Using a scaling argument similar to (3.15)–(3.16), one can show that on h^2-parallelograms and regular h^2-parallelepipeds,

$$\|\phi - \mathcal{Q}_h^{k-1}\phi\| \leq Ch^j \|\phi\|_j, \quad 1 \leq j \leq k. \qquad (3.18)$$

Moreover, the above bound holds with $j = 1$ on general quadrilaterals and hexahedra and with $j = 2$ on h^2-parallelepipeds.

Lemma 3.5. *For general quadrilaterals and hexahedra there exists a constant C independent of h such that for any finite element function φ*

$$\|\varphi\|_{j,E} \le Ch^{-1}\|\varphi\|_{j-1,E}, \quad j = 1, \ldots, k. \tag{3.19}$$

Proof. Let $\tilde{\varphi} = \varphi \circ F_E(\hat{x})$. Using (2.9), we have

$$|\varphi|_{1,E} \le \|DF_E^{-1}\|_{0,\infty,E}\|J_E\|_{0,\infty,\hat{E}}^{1/2}|\tilde{\varphi}|_{1,\hat{E}} \le C\|DF_E^{-1}\|_{0,\infty,E}\|J_E\|_{0,\infty,\hat{E}}^{1/2}\|\tilde{\varphi}\|_{\hat{E}}$$

$$\le C\|DF_E^{-1}\|_{0,\infty,E}\|J_E\|_{0,\infty,\hat{E}}^{1/2}\|J_{F_E}^{-1}\|_{0,\infty,E}^{1/2}\|\varphi\|_E \le Ch^{-1}h^{d/2}h^{-d/2}\|\varphi\|_E$$

$$\le Ch^{-1}\|\varphi\|_E.$$

The general case follows by applying the above bound to any derivative of φ. □

We will make use of the following continuity bounds for the mixed projection operators Π_*^k and Π_{RT}^k.

Lemma 3.6. *There exists a constant C independent of h such that on h^2-parallelograms and regular h^2-parallelepipeds*

$$\|\Pi_*^k\mathbf{q}\|_{j,E} \le C\|\mathbf{q}\|_{j,E}, \quad j = 1, \ldots, k+1, \tag{3.20}$$

$$\|\Pi_{RT}^{k-1}\mathbf{q}\|_{j,E} \le C\|\mathbf{q}\|_{j,E}, \quad j = 1, \ldots, k. \tag{3.21}$$

The above bounds also hold with $j = 1$ on h^2-parallelepipeds. Furthermore, on general quadrilaterals or hexahedra

$$\|\Pi_*^k\mathbf{q}\|_{\mathrm{div},E} + \|\Pi_{RT}^{k-1}\mathbf{q}\|_{\mathrm{div},E} \le C\|\mathbf{q}\|_{1,E}. \tag{3.22}$$

Proof. It follows from (3.12) and the triangle inequality that

$$\|\Pi_*^k\mathbf{q}\|_{0,E} \le \|\mathbf{q}\|_{1,E}.$$

Let $\mathcal{P}_E^j$ be the $L^2(E)$-projection onto $\mathcal{P}^j(E, \mathbb{R}^d)$. It is well known[19] that $\|\mathbf{q} - \mathcal{P}_E^j\mathbf{q}\|_E \le Ch^{j+1}\|\mathbf{q}\|_{j+1,E}$. Using (3.19), we have for any $j = 1, \ldots, k+1$,

$$|\Pi_*^k\mathbf{q}|_{j,E} = |\Pi_*^k\mathbf{q} - \mathcal{P}_E^{j-1}\mathbf{q}|_{j,E} \le Ch^{-j}\|\Pi_*^k\mathbf{q} - \mathcal{P}_E^{j-1}\mathbf{q}\|_{0,E}$$

$$\le Ch^{-j}(\|\Pi_*^k\mathbf{q} - \mathbf{q}\|_{0,E} + \|\mathbf{q} - \mathcal{P}_E^{j-1}\mathbf{q}\|_{0,E}) \le C\|\mathbf{q}\|_j,$$

where we also used (3.12), (3.13) and (3.18). This completes the proof of (3.20). The proof of (3.21) is similar. The proof of (3.22) uses a scaling argument similar to (3.15)–(3.16) for the divergence and a scaling argument using (3.8) for the L^2-norm. Details can be found in Lemma 3.6 in Ref. 27. □

Remark 3.1. For the rest of the paper, all results are stated for h^2-parallelograms and regular h^2-parallelepipeds. We note that the results also hold in 3d on h^2-parallelepipeds with $k = 1$, except for the pressure superconvergence.

In the next two lemmas we bound two terms arising in the error analysis due to the use of the quadrature rule. We use the notation $\varphi \in W^{k,\infty}_{\mathcal{T}_h}$ if $\varphi \in W^{k,\infty}(E) \ \forall E \in \mathcal{T}_h$ and $\|\varphi\|_{k,\infty,E}$ is uniformly bounded independently of h.

Lemma 3.7. *On h^2-parallelograms and regular h^2-parallelepipeds, if $K^{-1} \in W^{k,\infty}_{\mathcal{T}_h}$, then there exists a constant C independent of h such that for all $\mathbf{v} \in \mathbf{V}^k_h$,*

$$\left|\left(K^{-1}\Pi^k_*\mathbf{u}, \mathbf{v} - \Pi^{k-1}_{RT}\mathbf{v}\right)_Q\right| \le Ch^k\|\mathbf{u}\|_k\|\mathbf{v}\|. \tag{3.23}$$

Proof. Let $\hat{\mathcal{P}}^k$ be the $L^2(\hat{E})$-orthogonal projection onto $\mathcal{P}^k(\hat{E}, \mathbb{R}^d)$. For any element $E \in \mathcal{T}_h$, we have

$$\left(K^{-1}\Pi^k_*\mathbf{u}, \mathbf{v} - \Pi^{k-1}_{RT}\mathbf{v}\right)_{Q,E} = \left(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}}, \hat{\mathbf{v}} - \hat{\Pi}^{k-1}_{RT}\hat{\mathbf{v}}\right)_{Q,\hat{E}}$$

$$= \left(\hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}}), \hat{\mathbf{v}} - \hat{\Pi}^{k-1}_{RT}\hat{\mathbf{v}}\right)_{Q,\hat{E}}$$

$$+ \left(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}}), \hat{\mathbf{v}} - \hat{\Pi}^{k-1}_{RT}\hat{\mathbf{v}}\right)_{Q,\hat{E}}.$$

The first term on right is equal to zero due to (2.53). For the second term we use Bramble–Hilbert lemma:

$$\left|\left(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}}), \hat{\mathbf{v}} - \hat{\Pi}^{k-1}_{RT}\hat{\mathbf{v}}\right)_{Q,\hat{E}}\right| \le C|\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}}|_{k,\hat{E}}\|\hat{\mathbf{v}} - \hat{\Pi}^{k-1}_{RT}\hat{\mathbf{v}}\|_{0,\hat{E}}.$$

Using (3.10) and (3.8), we obtain

$$|\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}}|_{k,\hat{E}} \le C\sum_{i=0}^k |\mathcal{K}^{-1}|_{k-i,\infty,\hat{E}}|\hat{\Pi}^k_*\hat{\mathbf{u}}|_{i,\hat{E}}$$

$$\le C\sum_{i=0}^k h^{k-i-d+2}\|K^{-1}\|_{k-i,\infty,E}h^{i+(d-2)/2}\|\Pi^k_*\mathbf{u}\|_{i,E}$$

$$\le Ch^{k-d/2+1}\|K^{-1}\|_{k,\infty,E}\|\Pi^k_*\mathbf{u}\|_{k,E}.$$

Therefore, using (3.8), (3.20) and (2.41), we get

$$\left|\left(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_*\hat{\mathbf{u}}), \hat{\mathbf{v}} - \hat{\Pi}^{k-1}_{RT}\hat{\mathbf{v}}\right)_{Q,\hat{E}}\right|$$

$$\le Ch^{k-d/2+1}\|K^{-1}\|_{k,\infty,E}\|\mathbf{u}\|_{k,E}h^{(d-2)/2}\|\mathbf{v}\|_{0,E}$$

$$\le Ch^k\|K^{-1}\|_{k,\infty,E}\|\mathbf{u}\|_{k,E}\|\mathbf{v}\|_{0,E}.$$

The proof is completed by summing over all elements. □

Lemma 3.8. *On h^2-parallelograms and regular h^2-parallelepipeds, if $K^{-1} \in W^{k,\infty}_{\mathcal{T}_h}$, then there exists a constant C independent of mesh size such that for all $\mathbf{q} \in \mathbf{V}^k_h$ and $\mathbf{v} \in \mathbf{V}^{k-1}_{RT,h}$*

$$\left|\sigma\left(K^{-1}\mathbf{q}, \mathbf{v}\right)\right| \le C\sum_{E \in \mathcal{T}_h} h^k\|K^{-1}\|_{k,\infty,E}\|\mathbf{q}\|_{k,E}\|\mathbf{v}\|_E. \tag{3.24}$$

Proof. For each $E \in \mathcal{T}_h$ we have

$$\sigma_E\big(K^{-1}\mathbf{q}, \mathbf{v}\big) = \sigma_{\hat{E}}\big(\hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\mathbf{q}}), \hat{\mathbf{v}}\big) + \sigma_{\hat{E}}\big(\mathcal{K}^{-1}\hat{\mathbf{q}} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\mathbf{q}}), \hat{\mathbf{v}}\big).$$

The first term on the right is equal to zero, since the tensor-product Gauss–Lobatto quadrature rule is exact for polynomials of degree up to $2k-1$. Using the Bramble–Hilbert lemma, (3.10) and (3.8), we bound the second term as follows:

$$\big|\sigma_{\hat{E}}\big(\mathcal{K}^{-1}\hat{\mathbf{q}} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\mathbf{q}}), \hat{\mathbf{v}}\big)\big| \leq C|\mathcal{K}^{-1}\hat{\mathbf{q}}|_{k,\hat{E}}\|\hat{\mathbf{v}}\|_{\hat{E}}$$

$$\leq C\sum_{i=0}^{k}|\mathcal{K}^{-1}|_{k-i,\infty,\hat{E}}|\hat{\mathbf{q}}|_{i,\hat{E}}\|\hat{\mathbf{v}}\|_{\hat{E}}$$

$$\leq Ch^{k-d/2+1}\|K^{-1}\|_{k,\infty,E}\|\mathbf{q}\|_{k,E}h^{(d-2)/2}\|\mathbf{v}\|_E$$

$$\leq Ch^{k}\|K^{-1}\|_{k,\infty,E}\|\mathbf{q}\|_{k,E}\|\mathbf{v}\|_E.$$

Summing over all $E \in \mathcal{T}_h$, we obtain (3.24). $\qquad\square$

3.1. *Optimal convergence for the velocity*

We subtract the numerical method (2.63)–(2.64) from the variational formulation (2.4)–(2.5) to obtain the error equations:

$$\big(K^{-1}\mathbf{u}, \mathbf{v}\big) - \big(K^{-1}\mathbf{u}_h, \mathbf{v}\big)_Q - \big(p - p_h, \nabla \cdot \mathbf{v}\big)$$
$$= -\langle g - \mathcal{R}_h^{k-1}g, \mathbf{v} \cdot \mathbf{n}\rangle_{\Gamma_D}, \quad \mathbf{v} \in \mathbf{V}_h^k, \tag{3.25}$$

$$\big(\nabla \cdot (\mathbf{u} - \mathbf{u}_h), w\big) = 0, \quad w \in W_h^{k-1}. \tag{3.26}$$

Note that due to (2.35), it follows from (3.26) that

$$\nabla \cdot (\Pi_*^k \mathbf{u} - \mathbf{u}_h) = 0. \tag{3.27}$$

If we take $\mathbf{v} = \Pi_*^k \mathbf{u} - \mathbf{u}_h$ in (3.25), then

$$\big(K^{-1}\mathbf{u}, \Pi_*^k \mathbf{u} - \mathbf{u}_h\big) - \big(K^{-1}\mathbf{u}_h, \Pi_*^k \mathbf{u} - \mathbf{u}_h\big)_Q$$
$$+ \langle g - \mathcal{R}_h^{k-1}g, (\Pi_*^k \mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\rangle_{\Gamma_D} = 0. \tag{3.28}$$

Let $\mathbf{w} \equiv \Pi_*^k \mathbf{u} - \mathbf{u}_h$ then an algebraic manipulation of the above gives

$$\big(K^{-1}\mathbf{w}, \mathbf{w}\big)_Q = -\big(K^{-1}\mathbf{u}, \mathbf{w}\big) + \big(K^{-1}\Pi_*^k \mathbf{u}, \mathbf{w}\big)_Q - \langle g - \mathcal{R}_h^{k-1}g, \mathbf{w} \cdot \mathbf{n}\rangle_{\Gamma_D}.$$

Moreover, rewriting the right-hand side gives

$$\big(K^{-1}\mathbf{w}, \mathbf{w}\big)_Q = -\big(K^{-1}\mathbf{u}, \mathbf{w} - \Pi_{RT}^{k-1}\mathbf{w}\big) - \langle g - \mathcal{R}_h^{k-1}g, \mathbf{w} \cdot \mathbf{n}\rangle_{\Gamma_D}$$
$$- \big(K^{-1}\big(\mathbf{u} - \Pi_*^k \mathbf{u}\big), \Pi_{RT}^{k-1}\mathbf{w}\big) - \big(K^{-1}\Pi_*^k \mathbf{u}, \Pi_{RT}^{k-1}\mathbf{w}\big)$$
$$+ \big(K^{-1}\Pi_*^k \mathbf{u}, \Pi_{RT}^{k-1}\mathbf{w}\big)_Q + \big(K^{-1}\Pi_*^k \mathbf{u}, \mathbf{w} - \Pi_{RT}^{k-1}\mathbf{w}\big)_Q. \tag{3.29}$$

Testing (2.4) with $\mathbf{w} - \Pi_{RT}^{k-1}\mathbf{w}$ and using that $\nabla \cdot \mathbf{w} = \nabla \cdot \Pi_{RT}^{k-1}\mathbf{w} = 0$, see (3.27) and (2.40), we can rewrite the first two terms in (3.29) as

$$-\left(K^{-1}\mathbf{u}, \mathbf{w} - \Pi_{RT}^{k-1}\mathbf{w}\right) - \langle g - \mathcal{R}_h^{k-1}g, \mathbf{w} \cdot \mathbf{n}\rangle_{\Gamma_D}$$

$$= \langle g, (\mathbf{w} - \Pi_{RT}^{k-1}\mathbf{w}) \cdot \mathbf{n}\rangle_{\Gamma_D} - \langle g - \mathcal{R}_h^{k-1}g, \mathbf{w} \cdot \mathbf{n}\rangle_{\Gamma_D} = 0,$$

using that, due to (2.61)–(2.62), $\langle \mathcal{R}_h^{k-1}g, (\mathbf{w} - \Pi_{RT}^{k-1}\mathbf{w}) \cdot \mathbf{n}\rangle_{\Gamma_D} = 0$ and $\langle g - \mathcal{R}_h^{k-1}g, \Pi_{RT}^{k-1}\mathbf{w} \cdot \mathbf{n}\rangle_{\Gamma_D} = 0$. For the third term on the right in (3.29) we use (3.12) and (2.41) to get

$$|(K^{-1}(\mathbf{u} - \Pi_*^k\mathbf{u}), \Pi_{RT}^{k-1}\mathbf{w})| \le Ch^k\|K^{-1}\|_{0,\infty}\|\mathbf{u}\|_k\|\mathbf{w}\|.$$

To bound the fourth and fifth terms on the right in (3.29), we use (3.24), (3.20) and (2.41):

$$|-\left(K^{-1}\Pi_*^k\mathbf{u}, \Pi_{RT}^{k-1}\mathbf{w}\right) + \left(K^{-1}\Pi_*^k\mathbf{u}, \Pi_{RT}^{k-1}\mathbf{w}\right)_Q| = |\sigma(K^{-1}\Pi_*^k\mathbf{u}, \Pi_{RT}^{k-1}\mathbf{w})|$$

$$\le Ch^k\|K^{-1}\|_{k,\infty}\|\mathbf{u}\|_k\|\mathbf{w}\|.$$

For the last term on the right in (3.29) we use (3.23):

$$|\left(K^{-1}\Pi_*^k\mathbf{u}, \mathbf{w} - \Pi_{RT}^{k-1}\mathbf{w}\right)_Q| \le Ch^k\|K^{-1}\|_{k,\infty}\|\mathbf{u}\|_k\|\mathbf{w}\|.$$

Combining the above bounds, we obtain from (3.29) that

$$\left(K^{-1}(\Pi_*^k\mathbf{u} - \mathbf{u}_h), \Pi_*^k\mathbf{u} - \mathbf{u}_h\right)_Q \le Ch^k\|K^{-1}\|_{k,\infty}\|\mathbf{u}\|_k\|\Pi_*^k\mathbf{u} - \mathbf{u}_h\|, \tag{3.30}$$

implying that

$$\|\Pi_*^k\mathbf{u} - \mathbf{u}_h\| \le Ch^k\|K^{-1}\|_{k,\infty}\|\mathbf{u}\|_k. \tag{3.31}$$

Bounds (3.31) and (3.27), together with (3.12) and (3.14), result in the following theorem.

Theorem 3.1. *Assume that the partition $\mathcal{T}_h$ consists of h^2-parallelograms in 2d or regular h^2-parallelepipeds in 3d. If $K^{-1} \in W_{\mathcal{T}_h}^{k,\infty}$, for the velocity $\mathbf{u}_h$ of the MFMFE method (2.63)–(2.64), there exists a constant C independent of h such that*

$$\|\mathbf{u} - \mathbf{u}_h\| \le Ch^k\|\mathbf{u}\|_k, \tag{3.32}$$

$$\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\| \le Ch^k\|\nabla \cdot \mathbf{u}\|_k. \tag{3.33}$$

4. Error Estimates for the Pressure

In this section, we use a standard inf-sup argument to prove optimal convergence for the pressure. We also employ a duality argument to establish superconvergence for the pressure.

4.1. *Optimal convergence for the pressure*

Theorem 4.1. *Assume that the partition $\mathcal{T}_h$ consists of h^2-parallelograms in $2d$ or regular h^2-parallelepipeds in $3d$. If $K^{-1} \in W^{k,\infty}_{\mathcal{T}_h}$, then for the pressure p_h of the MFMFE method (2.63)–(2.64), there exists a constant C independent of h such that*

$$\|p - p_h\| \leq Ch^k(\|\mathbf{u}\|_k + \|p\|_k). \tag{4.1}$$

Proof. We first note that the RT_{k-1} spaces $\mathbf{V}^{k-1}_{RT,h} \times W^{k-1}_h$ on general quadrilaterals and hexahedra satisfy an inf-sup condition similar to (2.37). The proof is the same as the argument in Lemma 2.1. Hence, using (3.25) and (2.61), we obtain

$$\|\mathcal{Q}^{k-1}_h p - p_h\| \leq \frac{1}{\beta} \sup_{0 \neq \mathbf{v} \in V^{k-1}_{RT,h}} \frac{\left(\mathcal{Q}^{k-1}_h p - p_h, \nabla \cdot \mathbf{v}\right)}{\|\mathbf{v}\|_{\mathrm{div}}}$$

$$= \frac{1}{\beta} \sup_{0 \neq \mathbf{v} \in V^{k-1}_{RT,h}} \frac{\left(K^{-1}(\Pi^k_* \mathbf{u} - \mathbf{u}_h), \mathbf{v}\right)_Q - \left(K^{-1}(\Pi^k_* \mathbf{u} - \mathbf{u}), \mathbf{v}\right) + \sigma(K^{-1}\Pi^k_* \mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|_{\mathrm{div}}}$$

$$\leq \frac{C}{\beta} h^k \|K^{-1}\|_{k,\infty} \|\mathbf{u}\|_k,$$

where we used (3.31), (3.12), (3.24), and (3.20) in the last inequality. The result then follows from (3.18) and the triangle inequality. $\square$

4.2. *Superconvergence of the pressure*

In this subsection, we prove superconvergence of the pressure, i.e. we show that $\|\mathcal{Q}^{k-1}_h p - p_h\|$ is $\mathcal{O}(h^{k+1})$ for the MFMFE method of order k. We also apply local postprocessing to obtain an improved approximation $p^*_h \in W^k_h$ such that $\|p - p^*_h\|$ is $\mathcal{O}(h^{k+1})$.

The following bound on the quadrature error will be used in the superconvergence analysis.

Lemma 4.1. *On h^2-parallelograms and regular h^2-parallelepipeds, if $K^{-1} \in W^{k+1,\infty}_{\mathcal{T}_h}$, then for all $\mathbf{q} \in \mathbf{V}^k_h$ and $\mathbf{v} \in \mathbf{V}^0_{RT,h}$, there exists a positive constant C independent of h such that*

$$\left|\sigma\left(K^{-1}\mathbf{q}, \mathbf{v}\right)\right| \leq C \sum_{E \in \mathcal{T}_h} h^{k+1} \|K^{-1}\|_{k+1,\infty,E} \|\mathbf{q}\|_{k+1,E} \|\mathbf{v}\|_{1,E}. \tag{4.2}$$

Proof. For any element E we have $\sigma_E\left(K^{-1}\mathbf{q}, \mathbf{v}\right) = \hat{\sigma}_{\hat{E}}\left(\mathcal{K}^{-1}\hat{\mathbf{q}}, \hat{\mathbf{v}}\right)$. Since the quadrature rule is exact for polynomials of degree up to $2k - 1$ and $k \geq 1$, then it is exact for polynomials of degree up to k. An application of the Bramble–Hilbert

lemma implies

$$|\hat{\sigma}_{\hat{E}}(\mathcal{K}^{-1}\hat{\mathbf{q}}, \hat{\mathbf{v}})|$$

$$\leq C\left(\left[\sum_{i=0}^{k}|\mathcal{K}^{-1}|_{i,\infty,\hat{E}}|\hat{\mathbf{q}}|_{k-i,\hat{E}}\right]|\hat{\mathbf{v}}|_{1,\hat{E}} + \left[\sum_{i=0}^{k+1}|\mathcal{K}^{-1}|_{i,\infty,\hat{E}}|\hat{\mathbf{q}}|_{k+1-i,\hat{E}}\right]\|\hat{\mathbf{v}}\|_{\hat{E}}\right),$$

where we used that $\hat{\mathbf{v}}$ is linear. Using (3.8) and (3.10), we obtain

$$\sigma_E(K^{-1}\mathbf{q}, \mathbf{v}) \leq Ch^{k+1}\|K^{-1}\|_{k+1,\infty,E}\|\mathbf{q}\|_{k+1,E}\|\mathbf{v}\|_{1,E}.$$

Summation of over all elements completes the proof.   □

The following result establishes superconvergence of the pressure if the H^2-elliptic regularity which is defined below holds. Let ϕ be the solution of

$$-\nabla \cdot K\nabla\phi = -(\mathcal{Q}_h^{k-1}p - p_h) \quad \text{in } \Omega, \quad \phi = 0 \quad \text{on } \partial\Omega. \tag{4.3}$$

We say that this problem satisfies H^2-elliptic regularity if

$$\|K\nabla\phi\|_1 + \|\phi\|_2 \leq C\|\mathcal{Q}_h^{k-1}p - p_h\| \tag{4.4}$$

with constant C which may depend on K and Ω but is independent of ϕ. Some sufficient conditions for (4.4) can be found in Refs. 26 and 32. In the proof of the theorem below, we follow the argument in Ref. 20 with appropriate modification to deal with the quadrature terms.

Theorem 4.2. *Assume that the partition $\mathcal{T}_h$ consists of h^2-parallelograms in $2d$ or regular h^2-parallelepipeds in $3d$. Assume also that $K^{-1} \in W_{\mathcal{T}_h}^{k+1,\infty}$, and that the H^2-elliptic regularity (4.4) holds. Then, for the pressure p_h of the MFMFE method (2.63)–(2.64), there exists a constant C independent of h such that*

$$\|\mathcal{Q}_h^{k-1}p - p_h\| \leq Ch^{k+1}(\|\mathbf{u}\|_k + \|\nabla \cdot \mathbf{u}\|_k). \tag{4.5}$$

Proof. The proof makes use of a duality argument. Let ϕ be the solution of (4.3). Denoting $-K\nabla\phi$ by $\mathbf{u}^*$, $(\mathbf{u}^*, \phi)$ satisfy

$$\left(K^{-1}\mathbf{u}^*, \mathbf{v}\right) - \left(\phi, \nabla \cdot \mathbf{v}\right) = 0, \quad \mathbf{v} \in H(\text{div}; \Omega), \tag{4.6}$$

$$\left(\nabla \cdot \mathbf{u}^*, q\right) = -\left(\mathcal{Q}_h^{k-1}p - p_h, q\right), \quad q \in L^2(\Omega). \tag{4.7}$$

Taking $\mathbf{v} = \mathbf{u} - \mathbf{u}_h$, $q = -(\mathcal{Q}_h^{k-1}p - p_h)$ and adding the two equations gives

$$\left(K^{-1}\mathbf{u}^*, \mathbf{u} - \mathbf{u}_h\right) - \left(\phi, \nabla \cdot (\mathbf{u} - \mathbf{u}_h)\right) - \left(\nabla \cdot \mathbf{u}^*, \mathcal{Q}_h^{k-1}p - p_h\right) = \|\mathcal{Q}_h^{k-1}p - p_h\|^2. \tag{4.8}$$

Consider the discretization of (4.6)–(4.7) as in (2.63)–(2.64) and let $(\mathbf{u}_h^*, \phi_h^*)$ be the solution of the discrete problem. We now use the Galerkin orthogonality (3.25)–(3.26) with $\mathbf{v} = \Pi_{RT}^{k-1}\mathbf{u}_h^*$ and $w = \mathcal{Q}_h^{k-1}\phi$ to get

$$\left(K^{-1}\mathbf{u}, \Pi_{RT}^{k-1}\mathbf{u}_h^*\right) - \left(K^{-1}\mathbf{u}_h, \Pi_{RT}^{k-1}\mathbf{u}_h^*\right)_Q - \left(\mathcal{Q}_h^{k-1}p - p_h, \nabla \cdot \Pi_{RT}^{k-1}\mathbf{u}_h^*\right)$$

$$- \left(\nabla \cdot (\mathbf{u} - \mathbf{u}_h), \mathcal{Q}_h^{k-1}\phi\right) = 0, \tag{4.9}$$

where we used that $(p - \mathcal{Q}_h^{k-1}p, \nabla \cdot \Pi_{RT}^{k-1}\mathbf{u}_h^*) = 0$ due to (3.17) and $\langle g - \mathcal{R}_h^{k-1}g, \Pi_{RT}^{k-1}\mathbf{u}_h^* \cdot \mathbf{n}\rangle_{\Gamma_D} = 0$ due to (2.61). Subtracting (4.9) from (4.8) and using the symmetry of $(K^{-1}\cdot, \cdot)$ and $(K^{-1}\cdot, \cdot)_Q$ gives

$$\left(K^{-1}\left(\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \mathbf{u}\right) - \left(K^{-1}\mathbf{u}^*, \mathbf{u}_h\right) + \left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \mathbf{u}_h\right)_Q$$

$$- \left(\phi - \mathcal{Q}_h^{k-1}\phi, \nabla \cdot (\mathbf{u} - \mathbf{u}_h)\right) - \left(\nabla \cdot \left(\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \mathcal{Q}_h^{k-1}p - p_h\right)$$

$$= \|\mathcal{Q}_h^{k-1}p - p_h\|^2.$$

Since $\nabla \cdot \Pi_{RT}^{k-1}\mathbf{u}_h^* = \nabla \cdot \mathbf{u}_h^*$, and $\left(\nabla \cdot (\mathbf{u}^* - \mathbf{u}_h^*), q\right) = 0$ holds for all $q \in W_h^{k-1}$ from the definition of $\mathbf{u}_h^*$, the last term in the left-hand side vanishes. Therefore, we have

$$\left(K^{-1}\left(\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \mathbf{u} - \mathbf{u}_h\right) - \sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \mathbf{u}_h\right) - \left(\phi - \mathcal{Q}_h^{k-1}\phi, \nabla \cdot (\mathbf{u} - \mathbf{u}_h)\right)$$

$$= \|\mathcal{Q}_h^{k-1}p - p_h\|^2. \tag{4.10}$$

with $\sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \mathbf{u}_h\right) = \left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \mathbf{u}_h\right) - \left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \mathbf{u}_h\right)_Q$. Observe that the difference of (4.6) and its discrete counterpart gives

$$\left(K^{-1}\mathbf{u}^*, \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right) - \left(K^{-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)_Q = 0,$$

because $\nabla \cdot (\Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h) = 0$. From this we obtain

$$\sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \mathbf{u}_h\right)$$

$$= \sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u}\right) - \sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)$$

$$= \sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u}\right) - \left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)$$

$$+ \left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)_Q$$

$$= \sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u}\right) + \left(K^{-1}\left(\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)$$

$$- \left(K^{-1}\left(\mathbf{u}_h^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)_Q,$$

and we can rewrite (4.10) further as

$$\left(K^{-1}\left(\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \mathbf{u} - \Pi_{RT}^{k-1}\mathbf{u}\right) + \left(K^{-1}\left(\mathbf{u}_h^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)_Q$$

$$- \sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u}\right) - \left(\phi - \mathcal{Q}_h^{k-1}\phi, \nabla \cdot (\mathbf{u} - \mathbf{u}_h)\right) = \|\mathcal{Q}_h^{k-1}p - p_h\|^2. \tag{4.11}$$

We will show that the terms on left above can be bounded as follows:

$$\left|\left(K^{-1}\left(\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \mathbf{u} - \Pi_{RT}^{k-1}\mathbf{u}\right)\right| \leq Ch^{k+1}\|\mathcal{Q}_h^{k-1}p - p_h\|\|\mathbf{u}\|_k, \tag{4.12}$$

$$\left|\left(K^{-1}\left(\mathbf{u}_h^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\right), \Pi_{RT}^{k-1}\mathbf{u} - \mathbf{u}_h\right)_Q\right| \leq Ch^{k+1}\|\mathcal{Q}_h^{k-1}p - p_h\|\|\mathbf{u}\|_k, \tag{4.13}$$

$$\left|\sigma\left(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u}\right)\right| \leq Ch^{k+1}\|\mathcal{Q}_h^{k-1}p - p_h\|\|\mathbf{u}\|_k, \tag{4.14}$$

$$\left|\left(\phi - \mathcal{Q}_h^{k-1}\phi, \nabla \cdot (\mathbf{u} - \mathbf{u}_h)\right)\right| \leq Ch^{k+1}\|\mathcal{Q}_h^{k-1}p - p_h\|\|\nabla \cdot \mathbf{u}\|_k, \tag{4.15}$$

which, combined with (4.11), imply the statement of the theorem. For (4.12), we note that

$$\|\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}_h^*\| \leq \|\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}^*\| + \|\Pi_{RT}^{k-1}(\Pi_{RT}^{k-1}\mathbf{u}^* - \mathbf{u}_h^*)\|$$

$$\leq \|\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}^*\| + C\|\Pi_{RT}^{k-1}\mathbf{u}^* - \mathbf{u}_h^*\|$$

$$\leq \|\mathbf{u}^* - \Pi_{RT}^{k-1}\mathbf{u}^*\| + C(\|\Pi_{RT}^{k-1}\mathbf{u}^* - \mathbf{u}^*\| + \|\mathbf{u}^* - \mathbf{u}_h^*\|) \leq Ch\|\mathbf{u}^*\|_1, \quad (4.16)$$

where we used (2.41), (3.12), and a bound for the discretization error

$$\|\mathbf{u}^* - \mathbf{u}_h^*\| \leq Ch\|\mathbf{u}^*\|_1, \quad (4.17)$$

which is obtained in a manner similar to the velocity error estimate (3.32). Bound (4.12) follows from the use of the Cauchy–Schwarz inequality, (4.16), (3.12), and (4.4). Bound (4.13) is obtained in a similar way, by adding and subtracting $\mathbf{u}^*$ in the first component and $\mathbf{u}$ in the second component, and using (4.17), (4.16), (3.12), (3.32), and (4.4). Bound (4.14) follows from

$$|\sigma(K^{-1}\Pi_{RT}^{k-1}\mathbf{u}_h^*, \Pi_{RT}^{k-1}\mathbf{u})| \leq |\sigma(K^{-1}(\Pi_{RT}^{k-1}\mathbf{u}_h^* - \Pi_{RT}^0\mathbf{u}^*), \Pi_{RT}^{k-1}\mathbf{u})|$$

$$+ |\sigma(K^{-1}\Pi_{RT}^0\mathbf{u}^*, \Pi_{RT}^{k-1}\mathbf{u})|$$

$$\leq C(h^k\|\mathbf{u}\|_k\|\Pi_{RT}^{k-1}\mathbf{u}_h^* - \Pi_{RT}^0\mathbf{u}^*\| + h^{k+1}\|\mathbf{u}\|_k\|\mathbf{u}^*\|_1)$$

$$\leq Ch^{k+1}\|\mathcal{Q}_h^{k-1}p - p_h\|\|\mathbf{u}\|_k,$$

where we used (3.24), (4.2), (3.21), (4.16), (3.12), and (4.4). Finally, (4.15) follows from (3.18), (3.33), and (4.4). □

Using the above result we can easily show superconvergence of the pressure at the Gauss points. For an element E, let $\|\!|\cdot|\!\|_E$ denote the discrete $L^2(E)$-norm computed by mapping to the reference element $\hat{E}$ and applying the tensor-product Gauss quadrature rule with k points in each variable. It is easy to see that $\|\!|w|\!\|_E = \|w\|_E$ for $w \in W_h^{k-1}(E)$. Assuming continuous pressure $p|_E$, let $p^I|_E \in W_h^{k-1}(E)$ be the Lagrange interpolant of $p|_E$ at the k^d Gauss points. It is shown in Lemma 4.3 of Ref. 23 that

$$\|\mathcal{Q}_h^{k-1}p - p^I\| \leq Ch^{k+1}\|p\|_{k+1}. \quad (4.18)$$

We now have

$$\|\!|p - p_h|\!\| = \|\!|p^I - p_h|\!\| = \|p^I - p_h\| \leq \|p^I - \mathcal{Q}_h^{k-1}p\| + \|\mathcal{Q}_h^{k-1}p - p_h\|$$

$$\leq Ch^{k+1}(\|\mathbf{u}\|_k + \|\nabla \cdot \mathbf{u}\|_k + \|p\|_{k+1}),$$

using (4.18) and (4.5).

We next show that the above superconvergence result for $\|\mathcal{Q}_h^{k-1}p - p_h\|$ can be used to compute a higher order approximation to the pressure p in the $L^2(\Omega)$-norm, using a variant of the local postprocessing proposed in Ref. 39. The postprocessing

idea is also utilized for *a posteriori* error estimation (see e.g. work of Ref. 34). Let $\tilde{W}_h^k$ be the L^2-orthogonal complement of W_h^0 in W_h^k. We now define $p_h^* \in W_h^k$ by

$$\mathcal{Q}_h^0 p_h^* = \mathcal{Q}_h^0 p_h, \tag{4.19}$$

$$(\nabla p_h^*, \nabla q)_E = -(K^{-1}\mathbf{u}_h, \nabla q)_E, \quad q \in \tilde{W}_h^k(E), \quad \forall E \in \mathcal{T}_h. \tag{4.20}$$

Theorem 4.3. *Under the assumption of Theorem 4.2, there exists a constant C independent of h such that*

$$\|p - p_h^*\| \le Ch^{k+1}(\|\mathbf{u}\|_k + \|\nabla \cdot \mathbf{u}\|_k + \|p\|_{k+1}). \tag{4.21}$$

Proof. Let $\tilde{\mathcal{Q}}_h^k$ be the L^2 orthogonal projection onto $\tilde{W}_h^k$. By the triangle inequality it is enough to estimate $\|\mathcal{Q}_h^k p - p_h^*\|$. Let $\tilde{p}_h := p_h^* - \mathcal{Q}_h^0 p_h$. Considering the decomposition $\mathcal{Q}_h^k p - p_h^* = (\mathcal{Q}_h^0 p - \mathcal{Q}_h^0 p_h) + (\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)$, it is sufficient to estimate $\|\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h\|$ by Theorem 4.2. Recalling that $\nabla p = -K^{-1}\mathbf{u}$, we have

$$(\nabla_h(p - p_h^*), \nabla_h q) = -(K^{-1}(\mathbf{u} - \mathbf{u}_h), \nabla_h q), \quad \forall q \in \tilde{W}_h^k,$$

where ∇_h is the element-wise gradient. From $p - p_h^* = (p - \mathcal{Q}_h^k p) + (\mathcal{Q}_h^0 p - \mathcal{Q}_h^0 p_h) + (\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)$ and by taking $q = \tilde{\mathcal{Q}}_h^k p - \tilde{p}_h$ in the above equation, we get

$$\|\nabla_h(\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)\| \le \|\nabla_h(p - \mathcal{Q}_h^k p)\| + \|K^{-1}(\mathbf{u} - \mathbf{u}_h)\| \le Ch^k(\|p\|_k + \|\mathbf{u}\|_k),$$

where we used the Bramble–Hilbert lemma, an inverse estimate, and (3.32). Since W_h^0 is the space of element-wise constants on $\mathcal{T}_h$, $\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h$ is orthogonal to element-wise constants. Then the element-wise Friedrichs' inequality yields $\|\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h\|_E \le Ch_E\|\nabla_h(\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)\|_E$ for all $E \in \mathcal{T}_h$. The conclusion follows by combining this and the above inequality. □

Remark 4.1. Instead of the postprocessing (4.19)–(4.20), one may use the postprocessing defined in Ref. 39 and obtain a numerical pressure that is convergent of order $\mathcal{O}(h^{k+1})$. The error analysis is almost the same as the above.

5. Numerical Results

In this section, we present numerical experiments on quadrilateral and hexahedral grids that validate the theoretical results in the previous sections. The method has been implemented in the finite element library deal.II.[8] The code is available in the deal.II code gallery.[5] In the first example, we test the method on a sequence of meshes obtained by a uniform isotropic refinement of an initial quadrilateral partition of the unit square. The boundary conditions are chosen to be of Dirichlet type for simplicity. The test case is constructed with the full permeability tensor coefficient

$$K = \begin{pmatrix} (x+1)^2 + y^2 & \sin(xy) \\ \sin(xy) & (x+1)^2 \end{pmatrix},$$
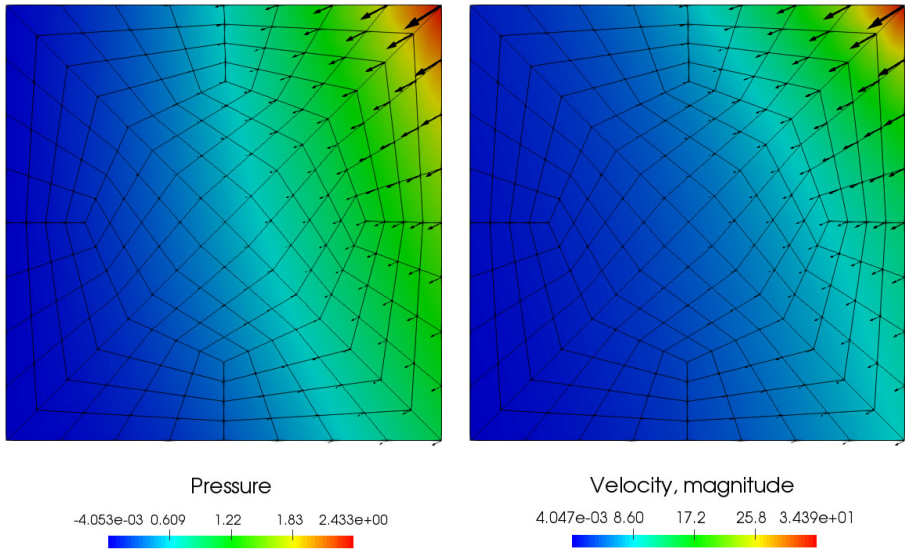
Fig. 2.   Computed solution for Example 1 on the third level of refinement.

and the analytical solution

$$p = x^3 y^4 + x^2 + \sin(xy)\cos(xy).$$

The computed pressure solution on the third level of refinement is shown in Fig. 2 (left), where the colors represent the pressure values and the arrows represent the velocity vectors. Similarly, Fig. 2 (right) shows the velocity solution, where colors represent the velocity magnitude. The numerical relative errors and convergence rates are obtained on a sequence of six mesh refinements and are reported in Table 1 for the MFMFE methods of order $k = 2, 3, 4$. We note that in all cases we see the predicted convergence rate of order $\mathcal{O}(h^k)$ for all variables in their natural norms, as well as superconvergence of the pressures at the Gauss points, i.e. $\|\|p - p_h\|\|$ is of order $\mathcal{O}(h^{k+1})$. We also observe $\mathcal{O}(h^{k+1})$ convergence for the postprocessed pressure. We note that the deterioration of the convergence rate of the divergence and the superconvergence rate of the pressure for the fourth-order method on the finest grid is due to the fact that these errors are very small and roundoff errors start having a noticeable effect.

In the second example, we focus on a 3d case. We let K be a full permeability tensor with variable coefficients

$$K = \begin{pmatrix} x^2 + (y+2)^2 & 0 & \cos(xy) \\ 0 & z^2 + 2 & \sin(xy) \\ \cos(xy) & \sin(xy) & (y+3)^2 \end{pmatrix},$$

Table 1. Relative errors and convergence rates for Example 1.

	$k = 2$									
	$\|\mathbf{u} - \mathbf{u}_h\|$		$\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|$		$\|p - p_h\|$		$\|\|p - p_h\|\|$		$\|p - p_h^*\|$	
h	Error	Rate	Error	Rate	Error	Rate	Error	Rate	Error	Rate
1/3	8.80E-02	—	1.46E-01	—	3.20E-02	—	5.80E-03	—	1.19E-02	—
1/6	2.36E-02	1.9	3.74E-02	2.0	7.90E-03	2.0	7.73E-04	2.9	1.42E-03	3.1
1/12	6.01E-03	2.0	9.41E-03	2.0	1.98E-03	2.0	1.18E-04	2.7	1.66E-04	3.1
1/24	1.50E-03	2.0	2.36E-03	2.0	4.96E-04	2.0	1.70E-05	2.8	1.94E-05	3.1
1/48	3.74E-04	2.0	5.89E-04	2.0	1.24E-04	2.0	2.30E-06	2.9	2.29E-06	3.1
1/96	9.31E-05	2.0	1.47E-04	2.0	3.10E-05	2.0	2.99E-07	2.9	2.78E-07	3.1
	$k = 3$									
	$\|\mathbf{u} - \mathbf{u}_h\|$		$\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|$		$\|p - p_h\|$		$\|\|p - p_h\|\|$		$\|p - p_h^*\|$	
h	Error	Rate	Error	Rate	Error	Rate	Error	Rate	Error	Rate
1/3	1.35E-02	—	1.96E-02	—	3.16E-03	—	4.36E-04	—	1.03E-03	—
1/6	1.69E-03	3.0	2.44E-03	3.0	3.95E-04	3.0	3.33E-05	3.7	5.33E-05	4.3
1/12	2.09E-04	3.0	3.04E-04	3.0	4.95E-05	3.0	2.48E-06	3.8	2.79E-06	4.3
1/24	2.59E-05	3.0	3.80E-05	3.0	6.19E-06	3.0	1.74E-07	3.8	1.55E-07	4.2
1/48	3.22E-06	3.0	4.75E-06	3.0	7.73E-07	3.0	1.17E-08	3.9	9.04E-09	4.1
1/96	4.02E-07	3.0	5.93E-07	3.0	9.67E-08	3.0	7.57E-10	4.0	5.44E-10	4.1
	$k = 4$									
	$\|\mathbf{u} - \mathbf{u}_h\|$		$\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|$		$\|p - p_h\|$		$\|\|p - p_h\|\|$		$\|p - p_h^*\|$	
h	Error	Rate	Error	Rate	Error	Rate	Error	Rate	Error	Rate
1/3	1.13E-03	—	1.52E-03	—	2.46E-04	—	2.83E-05	—	5.17E-05	—
1/6	6.84E-05	4.1	9.24E-05	4.0	1.52E-05	4.0	1.00E-06	4.8	1.26E-06	5.4
1/12	4.20E-06	4.0	5.74E-06	4.0	9.50E-07	4.0	3.55E-08	4.8	3.20E-08	5.3
1/24	2.59E-07	4.0	3.58E-07	4.0	5.94E-08	4.0	1.20E-09	4.9	8.74E-10	5.2
1/48	1.61E-08	4.0	2.25E-08	4.0	3.71E-09	4.0	3.98E-11	4.9	2.59E-11	5.1
1/96	1.00E-09	4.0	4.96E-09	2.2	2.32E-10	4.0	8.78E-12	2.2	8.72E-12	1.6

and solve the problem with Dirichlet boundary conditions and the analytical pressure solution chosen as follows

$$p = x^4 y^3 + x^2 + yz^2 + \cos(xy) + \sin(z).$$

The initial computational domain is obtained as a smooth map of the unit cube, i.e. we start with a $4 \times 4 \times 4$ unit cube mesh and then apply the following transformation to its points

$$x = \hat{x} + 0.03 \cos(3\pi\hat{x}) \cos(3\pi\hat{y}) \cos(3\pi\hat{z}),$$

$$y = \hat{y} - 0.04 \cos(3\pi\hat{x}) \cos(3\pi\hat{y}) \cos(3\pi\hat{z}),$$

$$z = \hat{z} + 0.05 \cos(3\pi\hat{x}) \cos(3\pi\hat{y}) \cos(3\pi\hat{z}).$$

The sequence of meshes on which we perform the convergence study is then obtained by a series of uniform refinements of the initial grid, described above. Figure 3 (left) presents the pressure solution, computed on the third level of refinement, where the colors represent the pressure values and the arrows depict the velocity vectors. The velocity magnitude is also shown in Fig. 3 (right). The computed numerical errors and convergence rates shown in Table 2 once again confirm the theoretical results
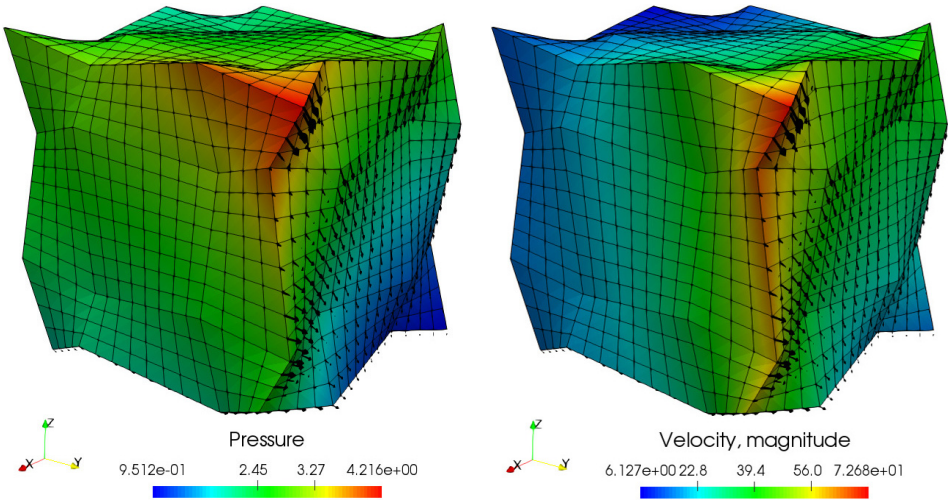
Fig. 3.   Computed solution for Example 2 on the third level of refinement.

Table 2.   Relative errors and convergence rates for Example 2.

	$\|\mathbf{u} - \mathbf{u}_h\|$		$\|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|$		$\|p - p_h\|$		$\|\|p - p_h\|\|$		$\|p - p_h^*\|$	
					$k = 2$					
h	Error	Rate	Error	Rate	Error	Rate	Error	Rate	Error	Rate
1/4	7.47E-03	—	2.92E-02	—	4.97E-03	—	1.63E-04	—	3.34E-04	—
1/8	1.82E-03	2.0	7.24E-03	2.0	1.24E-03	2.0	2.23E-05	2.9	3.99E-05	3.1
1/16	4.51E-04	2.0	1.81E-03	2.0	3.11E-04	2.0	3.07E-06	2.9	4.86E-06	3.0
1/32	1.12E-04	2.0	4.51E-04	2.0	7.77E-05	2.0	4.12E-07	2.9	6.00E-07	3.0
1/64	2.80E-05	2.0	1.13E-04	2.0	1.94E-05	2.0	5.38E-08	2.9	7.47E-08	3.0
					$k = 3$					
h	Error	Rate	Error	Rate	Error	Rate	Error	Rate	Error	Rate
1/4	5.06E-04	—	2.01E-03	—	2.03E-04	—	3.78E-06	—	1.23E-05	—
1/8	6.37E-05	3.0	2.46E-04	3.0	2.54E-05	3.0	2.56E-07	3.9	6.93E-07	4.2
1/16	7.93E-06	3.0	3.05E-05	3.0	3.17E-06	3.0	1.87E-08	3.8	4.06E-08	4.1
1/32	9.87E-07	3.0	3.81E-06	3.0	3.97E-07	3.0	1.35E-09	3.8	2.46E-09	4.0
1/64	1.21E-07	3.0	4.88E-07	3.0	4.96E-08	3.0	8.83E-11	3.9	1.50E-10	4.0

from the error analysis section. We see the optimal $\mathcal{O}(h^k)$ order of convergence for all variables, and also $\mathcal{O}(h^{k+1})$ superconvergence for the pressure.

In summary, the numerical experiments confirm the theoretical convergence results for the higher order MFMFE method both on h^2-parallelograms and regular h^2-parallelepipeds.

Acknowledgments

References

1. I. Aavatsmark, An introduction to multipoint flux approximations for quadrilateral grids, *Comput. Geosci.* **6** (2002) 405–432.

2. I. Aavatsmark, T. Barkve, O. Bøe and T. Mannseth, Discretization on unstructured grids for inhomogeneous, anisotropic media I. Derivation of the methods, *SIAM J. Sci. Comput.* **19** (1998) 1700–1716.

3. I. Aavatsmark, G. T. Eigestad, R. A. Klausen, M. F. Wheeler and I. Yotov, Convergence of a symmetric MPFA method on quadrilateral grids, *Comput. Geosci.* **11** (2007) 333–345.

4. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series, Vol. 55 (U.S. Government Printing Office, 1964), For sale by the Superintendent of Documents.

5. I. Ambartsumyan and E. Khattatov, Higher order multipoint flux mixed finite element (MFMFE) methods code gallery program, https://dealii.org/developer/doxygen/deal.II/code_gallery_MultipointFluxMixedFiniteElementMethods.html.

6. T. Arbogast, C. N. Dawson, P. T. Keenan, M. F. Wheeler and I. Yotov, Enhanced cell-centered finite differences for elliptic equations on general geometry, *SIAM J. Sci. Comput.* **19** (1998) 404–425.

7. T. Arbogast, M. F. Wheeler and I. Yotov, Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences, *SIAM J. Numer. Anal.* **34** (1997) 828–852.

8. D. Arndt, W. Bangerth, D. Davydov, T. Heister, L. Heltai, M. Kronbichler, M. Maier, J.-P. Pelteret, B. Turcksin and D. Wells, The deal.II library, version 8.5, *J. Numer. Math.* **25** (2017) 137–146.

9. D. N. Arnold, D. Boffi and R. S. Falk, Quadrilateral $H(\mathrm{div})$ finite elements, *SIAM J. Numer. Anal.* **42** (2005) 2429–2451.

10. D. N. Arnold and F. Brezzi, Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates, *RAIRO Modél. Math. Anal. Numér.* **19** (1985) 7–32.

11. D. N. Arnold, R. S. Falk and R. Winther, Finite element exterior calculus, homological techniques, and applications, *Acta Numer.* **15** (2006) 1–155.

12. D. N. Arnold, R. S. Falk and R. Winther, Finite element exterior calculus: From Hodge theory to numerical stability, *Bull. Amer. Math. Soc.* (*N.S.*) **47** (2010) 281–354.

13. J. Baranger, J.-F. Maitre and F. Oudin, Connection between finite volume and mixed finite element methods, *RAIRO Modél. Math. Anal. Numér.* **30** (1996) 445–465.

14. F. Brezzi, J. Douglas, Jr., R. Duràn and M. Fortin, Mixed finite elements for second order elliptic problems in three variables, *Numer. Math.* **51** (1987) 237–250.

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

16. F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Series in Computational Mathematics, Vol. 15 (Springer-Verlag, 1991).

17. F. Brezzi, M. Fortin and L. D. Marini, Error analysis of piecewise constant pressure approximations of Darcy's law, *Comput. Methods Appl. Mech. Eng.* **195** (2006) 1547–1559.

18. Z. Cai, J. Douglas, Jr. and M. Park, Development and analysis of higher order finite volume methods over rectangles for elliptic equations, *Adv. Comput. Math.* **19** (2003) 3–33.

19. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics, Vol. 40 (Society for Industrial and Applied Mathematics (SIAM), 2002), reprint of the 1978 original (North-Holland, MR0520174 (58 #25001)).

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

21. M. G. Edwards, Unstructured, control-volume distributed, full-tensor finite-volume schemes with flow based grids, *Comput. Geosci.* **6** (2002) 433–452.

22. M. G. Edwards and C. F. Rogers, Finite volume discretization with imposed flux continuity for the general tensor pressure equation, *Comput. Geosci.* **2** (1999) 259–290.

23. R. E. Ewing, R. D. Lazarov and J. Wang, Superconvergence of the velocity along the Gauss lines in mixed finite element methods, *SIAM J. Numer. Anal.* **28** (1991) 1015–1029.

24. R. E. Ewing, M. M. Liu and J. Wang, Superconvergence of mixed finite element approximations over quadrilaterals, *SIAM J. Numer. Anal.* **36** (1999) 772–787.

25. G. P. Galdi, *An Introduction to the Mathematical Theory of the Navier–Stokes Equations*, 2nd edn., Springer Monographs in Mathematics (Springer, 2011).

26. P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monographs and Studies in Mathematics, Vol. 24 (Pitman Advanced Publishing Program, 1985).

27. R. Ingram, M. F. Wheeler and I. Yotov, A multipoint flux mixed finite element method on hexahedra, *SIAM J. Numer. Anal.* **48** (2010) 1281–1312.

28. R. A. Klausen and A. F. Stephansen, Convergence of multi-point flux approximations on general grids and media, *Int. J. Numer. Anal. Model.* **9** (2012) 584–606.

29. R. A. Klausen and R. Winther, Convergence of multipoint flux approximations on quadrilateral grids, *Numer. Methods Partial Differential Equations* **22** (2006) 1438–1454.

30. R. A. Klausen and R. Winther, Robust convergence of multi point flux approximation on rough grids, *Numer. Math.* **104** (2006) 317–337.

31. J. J. Lee and R. Winther, Local coderivatives and approximation of Hodge Laplace problems, *Math. Comp.* **87** (2018) 2709–2735.

32. J.-L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Vol. I (Springer-Verlag, 1972), Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

33. K. Lipnikov, M. Shashkov and I. Yotov, Local flux mimetic finite difference methods, *Numer. Math.* **112** (2009) 115–152.

34. C. Lovadina and R. Stenberg, Energy norm *a posteriori* error estimates for mixed finite element methods, *Math. Comp.* **75** (2006) 1659–1674 (electronic).

35. J.-C. Nédélec, A new family of mixed finite elements in $\mathbf{R}^3$, *Numer. Math.* **50** (1986) 57–81.

36. P.-A. Raviart and J. M. Thomas, *A Mixed Finite Element Method for 2nd Order Elliptic Problems*, Lecture Notes in Mathematics, Vol. 606 (Springer-Verlag, 1977), pp. 292–315.

37. J. E. Roberts and J.-M. Thomas, Mixed and hybrid methods, in *Handbook of Numerical Analysis*, Vol. II (North-Holland, 1991), pp. 523–639.

38. T. F. Russell and M. F. Wheeler, Finite element and finite difference methods for continuous flows in porous media, *Math. Reser. Simul.* **1** (1983) 35–106.

39. R. Stenberg, Postprocessing schemes for some mixed finite elements, *RAIRO Modél. Math. Anal. Numér.* **25** (1991) 151–167.

40. A. F. Stephansen, Convergence of the multipoint flux approximation *L*-method on general grids, *SIAM J. Numer. Anal.* **50** (2012) 3163–3187.

41. M. Vohralík, Equivalence between lowest-order mixed finite element and multi-point finite volume methods on simplicial meshes, *Math. Model. Numer. Anal.* **40** (2006) 367–391.

42. M. Vohralík and B. I. Wohlmuth, Mixed finite element methods: Implementation with one unknown per element, local flux expressions, positivity, polygonal meshes, and relations to other methods, *Math. Models Methods Appl. Sci.* **23** (2013) 803–838.

43. J. Wang and T. Mathew, Mixed finite element methods over quadrilaterals, in *Conference on Advances in Numerical Methods and Applications*, eds. I. T. Dimov, B. Sendov and P. Vassilevski (World Scientific, 1994), pp. 203–214.

44. A. Weiser and M. F. Wheeler, On convergence of block-centered finite differences for elliptic problems, *SIAM J. Numer. Anal.* **25** (1988) 351–375.

45. M. Wheeler, G. Xue and I. Yotov, A multipoint flux mixed finite element method on distorted quadrilaterals and hexahedra, *Numer. Math.* **121** (2012) 165–204.

46. M. F. Wheeler, G. Xue and I. Yotov, A multiscale mortar multipoint flux mixed finite element method, *ESAIM Math. Model. Numer. Anal.* **46** (2012) 759–796.

47. M. F. Wheeler and I. Yotov, A multipoint flux mixed finite element method, *SIAM J. Numer. Anal.* **44** (2006) 2082–2106.

48. A. Younes and V. Fontaine, Hybrid and multi-point formulations of the lowest-order mixed methods for Darcy's flow on triangles, *Internat. J. Numer. Methods Fluids* **58** (2008) 1041–1062.