

# ICES REPORT 15-17

---

July 2015

## A Generalized Mimetic Finite Difference Method and Two-Point Flux Schemes over Voronoi Diagrams

by

Omar Al-Hinai, Mary F. Wheeler, Ivan Yotov



**The Institute for Computational Engineering and Sciences**  
The University of Texas at Austin  
Austin, Texas 78712

*Reference: Omar Al-Hinai, Mary F. Wheeler, Ivan Yotov, "A Generalized Mimetic Finite Difference Method and Two-Point Flux Schemes over Voronoi Diagrams," ICES REPORT 15-17, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, July 2015.*

# A Generalized Mimetic Finite Difference Method and Two-Point Flux Schemes over Voronoi Diagrams

Omar Al-Hinai\* and Mary F. Wheeler†

*Center for Subsurface Modeling,  
Institute for Computational Engineering and Sciences,  
The University of Texas at Austin, 201 East 24th Street,  
Austin, TX 78712, USA.*

\*ohinai@ices.utexas.edu

†mfw@ices.utexas.edu

Ivan Yotov

*Department of Mathematics,  
University of Pittsburgh,  
301 Thackeray Hall,  
Pittsburgh, PA 15260, USA.*

yotov@math.pitt.edu

June 24, 2015

## Abstract

We propose a generalization of the Mimetic Finite Difference (MFD) method that extends the family of convergent schemes to include Two-Point Flux Approximations (TPFA) over general Voronoi meshes. We prove first-order convergence in pressure and flux, and superconvergence of the pressure under further restrictions. We construct  $H(\Omega; \text{div})$  basis functions that correspond to the generalized form over the square element. We present numerical results that support the theory.

## 1 Introduction

Finite Volume Methods (FVM) are among the most widely used techniques for the numerical solution of partial differential equations, especially those governing fluid flow problems [23]. The popularity of FVM is due to the fact that they are locally conservative by construction, fast and robust for a wide range of problems, and can preserve the maximum principle. FVM can be defined over a wide variety of mesh types, including triangular/tetrahedral, quadrilateral/hexahedral and Voronoi type meshes [5, 40, 23]. The reader is referred to [21, 20] for work

on unstructured grids and full tensor permeabilities, and [1] for distorted grids and multi-point flux approximations (MPFA).

FVM can be related to other discretization methods such as the Finite Difference Method and the Finite Element Method. Of note is the connection established in [42] for Mixed Finite Elements (MFE), where the lowest order Raviart-Thomas element ( $RT_0$ ) [41] over rectangles was shown to be equivalent to cell-centered finite differences. The technique demonstrated how to reduce the number of degrees of freedom for  $RT_0$  and produce a linear symmetric positive definite system for the pressure variable. This connection enabled a proof of convergence for FVM based on the theory of MFE [48]. Many authors have subsequently built on these ideas, producing further connections and more theoretical results. An extension for full-tensor coefficients was developed in [3, 4]. Superconvergence for both the scalar and vector unknowns has been shown in [39, 18, 22]. Further extensions include the use of the lowest order Brezzi-Douglas-Marini element ( $BDM_1$ ) [11] for connecting MFE with MPFA [50, 28, 49]. A similar connection based on broken  $RT_0$  elements was developed in [31, 32].

While there has been success in demonstrating the connection between MFE and FVM, some types of FVM have not been connected with MFE. An example is the two-point flux approximation (TPFA) over Voronoi diagrams and K-orthogonal grids (grids aligned with the eigenvectors of the permeability tensor) [23]. Connecting such a method to MFE would require the construction of discrete  $H(\Omega; \text{div})$  spaces over convex polyhedra with an arbitrary number of faces. While such spaces have been constructed via a local triangulation of the cells [33], it remains uncertain if such an approach can be reduced to TPFA. The work done in [13, 14] on the Mimetic Finite Difference (MFD) method opens up some possibilities. The authors of [13] use theoretical tools from MFE to demonstrate convergence of the MFD method over a very general set of polyhedral cells. This is accomplished by avoiding the explicit construction of the velocity variable on the interior, and directly solving for flux degrees of freedom on the faces of the cells. The definition of the MFD method in [14] allowed for a reduction to TPFA for the case of regular polyhedra and rectangles. Later work in [19] extended the MFD method to include TPFA for acute triangular cells. The authors of [37] provided generalizations of the MFD method that included TPFA over centroidal Voronoi diagrams as well as MPFA methods over general polyhedra; see also related work in [30].

The objective of our work is to provide a further generalization of the MFD method that allows for TPFA type methods over general Voronoi diagrams. The reduction to TPFA results in a faster solution to the MFD method as well as solutions that satisfy the maximum principle. We propose a modified form of the consistency condition ( $S2$ ) in [13], which we call here ( $\widetilde{S2}$ ). The modification is inspired by the consistency conditions proposed by [19] and [37]. As a consequence, the method also relates the MFD method to the point-centered FVM. This is accomplished without the use of dual grids or local triangulations of the cells. Section 2 defines the new generalization of the MFD method. The stability of the new method is shown in Section 2.1, which is followed by first-order convergence proofs for the velocity in Section 2.2 and pressure variables in Section 2.3. Conditions for superconvergence of the method are presented in Section 2.4, which includes superconvergence with special quadrature in Section 2.4.1. Novel  $H(\Omega; \text{div})$  shape functions that satisfy the superconvergence criteria are presented in Section 2.4.2. We then discuss matrix construction in Section 3, and the relation to TPFA schemes on Voronoi grids in Section 4. Numerical results confirming the theory are demonstrated in Section 5.

The MFD method has found widespread application in many areas of numerical modeling. The central theme is to create discretizations that mimic the fundamental properties of the underlying partial differential equation. The MFD method has been used to solve elliptic and parabolic equations and has found applications in the areas of gas dynamics [45], Maxwell's equations [27, 9] and Stokes flow [16]. It is also possible to apply the MFD method to cell-faced [13] and nodal discretizations [10]. The MFD method can be used both to define higher-order methods [47], and on meshes with curved faces [36]. More recent work in higher order methods is in the related area of the Virtual Element Method [15]. Earlier examples of the MFD method referred to as the Support Operator Method (SOM) can be found in [44, 45, 26].

## 2 A Generalized MFD Method

Consider the homogeneous Dirichlet problem written in mixed form,

$$\begin{aligned} u &= -K\nabla p && \text{in } \Omega, \\ \nabla \cdot u &= f && \text{in } \Omega, \\ p &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{2.1}$$

In porous media flow,  $u$  is Darcy velocity,  $p$  is pressure and  $K$  is the permeability tensor. The domain  $\Omega \subset \mathbb{R}^d$  ( $d = 2, 3$ ) is assumed to be polyhedron with Lipschitz boundaries and is divided into a non-overlapping, conformal partition  $\mathcal{T}_h$  of elements  $E$ . In three dimensions,  $E$  is a polyhedron with planar polygonal faces. In two dimensions,  $E$  is a polygon. For the remainder of this work, we refer to elements  $E$  as polyhedra, with the understanding that the results also apply directly to the two-dimensional case. We make the following standard assumptions for the mesh used in previous work [13, 37]:

1. Elements and faces are shape-regular and non-degenerate;
2. Elements are star-shaped polyhedra;
3. Faces are star-shaped polygons.

We define the following mesh related quantities:

1.  $N_Q$  is the total number of elements in the mesh;
2.  $N_X$  is the total number of faces in the mesh;
3.  $|E|$  is the volume of  $E$ ;
4.  $h_E$  is the diameter of  $E$ ;
5.  $h = \max_{E \in \mathcal{T}_h} h_E$ ;
6.  $|e|$  is the area of face  $e$ ;
7.  $n^e$  is a fixed unit normal on face  $e$ ;

8.  $n_E^e$  is the unit normal on face  $e$  pointing out of element  $E$ ;
9.  $k_E$  is the number of faces of element  $E$ .

We note that the mesh assumptions imply that

$$|E| \sim h_E^d \text{ and } |e| \sim h_E^{d-1}. \quad (2.2)$$

Let the space of discrete pressures  $Q_h \subset \mathbb{R}^{N_Q}$  and fluxes  $X_h \subset \mathbb{R}^{N_X}$  be defined, respectively, as

$$\begin{aligned} Q_h &= \{q_h = \{q_E\}_{E \in \mathcal{T}_h} \text{ such that } q_E \in \mathbb{R}\}, \\ X_h &= \{v_h = \{v_E^e\}_{E \in \mathcal{T}_h}^{e \in \partial E} \text{ such that } v_E^e \in \mathbb{R}\}. \end{aligned}$$

Each pressure degree of freedom ( $q_E$ ) represents the pressure on element  $E$ . Each velocity degree of freedom ( $v_E^e$ ) represents the flux normal to the face  $e$  facing out of element  $E$ . The above definition for  $X_h$  results in two degrees of freedom per internal face in the mesh and a single degree of freedom for each boundary face. We require continuity of velocity in  $X_h$ , that is, for two adjacent elements  $E_a$  and  $E_b$  sharing a face  $e$ , let

$$v_{E_a}^e + v_{E_b}^e = 0.$$

The continuity of flux reduces the total number of degrees of freedom in  $X_h$  to  $N_X$ .

In the following, we make use of the usual notation for Lebesgue spaces  $L^p(\Omega)$ , Sobolev spaces  $W^{k,p}(\Omega)$  and Hilbert spaces  $H^k(\Omega)$ , see [8]. For  $q \in L^1(\Omega)$ , define a projection operator on  $Q_h$  as

$$(q^J)_E = \frac{1}{|E|} \int_E q \, dV. \quad (2.3)$$

For  $v \in (L^s(\Omega))^d$ ,  $s > 2$  and  $\nabla \cdot v \in L^2(\Omega)$ , define the projection operator on  $X_h$  as

$$(v^J)_E^e = \frac{1}{|e|} \int_e v \cdot n_E^e \, dS. \quad (2.4)$$

Let the permeability tensor  $K$  be symmetric and positive definite, and let

$$K_{ij} \in W^{1,\infty}(\Omega). \quad (2.5)$$

Let  $\bar{K}_E$  denote a constant tensor over  $E$  such that

$$\max_{ij} \|K_{ij} - \bar{K}_{E,ij}\|_{L^\infty(E)} \leq Ch_E. \quad (2.6)$$

Throughout the paper,  $C$  denotes a generic positive constant independent of  $h$ . Since  $K$  is positive definite, we also have that

$$\max_{ij} \|K_{ij}^{-1} - \bar{K}_{E,ij}^{-1}\|_{L^\infty(E)} \leq Ch_E. \quad (2.7)$$

Following the standard steps of the MFD method [34], we define the discrete divergence operator  $\mathcal{DIV} : X_h \rightarrow Q_h$  as

$$(\mathcal{DIV} v_h)_E = \mathcal{DIV} v_E := \frac{1}{|E|} \sum_{e \in \partial E} v_E^e |e|. \quad (2.8)$$

The proof of the next lemma follow from (2.3), (2.4), (2.8), and the divergence theorem.

**Lemma 1.** For  $v \in (L^s(\Omega))^d$ ,  $s > 2$  and  $\nabla \cdot v \in L^2(\Omega)$ ,

$$(\mathcal{D}\mathcal{I}\mathcal{V}v^I)_E = (\nabla \cdot v)_E^I. \quad (2.9)$$

Next, define the scalar inner product as

$$[p_h, q_h]_{Q_h} := \sum_{E \in \mathcal{T}_h} p_E q_E |E|.$$

The flux bilinear form is defined as

$$[u_h, v_h]_{X_h} := \sum_{E \in \mathcal{T}_h} [u_E, v_E]_E = \sum_{E \in \mathcal{T}_h} v_E^T \mathbf{M}_E u_E, \quad (2.10)$$

where  $u_E \in \mathbb{R}^{k_E}$  is the vector with components  $u_E^e$  and  $\mathbf{M}_E \in \mathbb{R}^{k_E \times k_E}$  is a given matrix. We discuss the construction of  $\mathbf{M}_E$  in Section 3. This form approximates the continuous velocity bilinear form, that is,

$$[u_h, v_h]_{X_h} \approx \int_{\Omega} K^{-1} u \cdot v \, dV.$$

We assume that  $[\cdot, \cdot]_{X_h}$  satisfies:

(S1) (*Stability*) There exist two positive constants,  $s_*$  and  $S^*$ , such that, for all  $E \in \mathcal{T}_h$  and  $\xi \in \mathbb{R}^{k_E}$ ,

$$s_* |E| \xi^T \xi \leq \xi^T \mathbf{M}_E \xi \leq S^* |E| \xi^T \xi, \quad (2.11)$$

and

$$\xi^T \mathbf{M}_E^T \mathbf{M}_E \xi \leq (S^*)^2 |E|^2 \xi^T \xi. \quad (2.12)$$

( $\widetilde{S}2$ ) (*Consistency*) For every element  $E \in \mathcal{T}_h$ , for every linear function  $q^1$  on  $E$ , and for every  $v_E \in X_h(E)$ ,

$$[(\bar{K}_E \nabla q^1)_E^I, v_E]_E = \sum_{e \in \partial E} v_E^e \int_e w_e q^1 \, dS - \int_E w_E q^1 \mathcal{D}\mathcal{I}\mathcal{V} v_E \, dV, \quad (2.13)$$

where the function  $w_E : E \rightarrow \mathbb{R}$  satisfies

$$\begin{aligned} \int_E w_E \, dV &= |E|, \\ \int_E g w_E \, dV &= |E| x_E, \end{aligned}$$

where  $x_E$  is a point in  $E$  and  $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the linear function

$$g(x, y, z) = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

and where the function  $w_e : e \rightarrow \mathbb{R}$  satisfies,

$$\int_e w_e dS = |e|, \quad (2.14)$$

$$\int_e g w_e dS = |e| x_e, \quad (2.15)$$

where  $x_e$  is a point in the plane of  $e$  and  $g$  is the linear function defined on Cartesian coordinates  $(x, y)$  in the plane of  $e$  as

$$g(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}.$$

The point  $x_e$  may lie outside of  $e$ .

Note that if  $x_E = C_E$ , the centroid of  $E$ , then  $w_E = 1$ . Also, if  $x_e = C_e$ , the centroid of  $e$ , then  $w_e = 1$ . We require that the functions  $w_e$  and  $w_E$  be bounded by positive constants independent of the mesh,

$$\begin{aligned} \max_{E \in \mathcal{T}_h} \|w_E\|_{L^\infty(E)} &\leq w_E^{\max}, \\ \max_{E \in \mathcal{T}_h} \max_{e \in \partial E} \|w_e\|_{L^\infty(e)} &\leq w_e^{\max}. \end{aligned}$$

**Remark.** *The use of weighting functions in  $(\widetilde{S2})$  is the only departure from the original  $(S2)$  condition in [13]. The addition of weighting functions to the integrals was inspired by the modification suggested in [19] and [37].*

Define a discrete gradient operator  $\mathcal{G} : Q_h \rightarrow X_h$  as the adjoint of  $\mathcal{DIV}$ ,

$$[\mathcal{G} p_h, u_h]_{X_h} = -[p_h, \mathcal{DIV} u_h]_{Q_h}. \quad (2.16)$$

Note that, due to the lack of a boundary term in (2.16), it imposes weakly the boundary condition  $p_h = 0$  on  $\partial\Omega$ .

The numerical method can be represented in a “weak” saddle-point form. That is, find  $u_h \in X_h$  and  $p_h \in Q_h$  such that

$$[u_h, v_h]_{X_h} - [p_h, \mathcal{DIV} v_h]_{Q_h} = 0, \quad \forall v_h \in X_h, \quad (2.17)$$

$$[\mathcal{DIV} u_h, q_h]_{Q_h} = [f^I, q_h]_{Q_h}, \quad \forall q_h \in Q_h. \quad (2.18)$$

**Remark.** *The theoretical results presented here follow the exposition found in [37]. We have made modifications when necessary to adapt to the modified consistency condition  $(\widetilde{S2})$ . Since the stability analysis does not depend on the consistency condition  $(\widetilde{S2})$ , the stability results from [13] and [37] still hold.*

## 2.1 Stability Analysis

The analysis follows the classical approach to stability of mixed methods. We start by defining the following norms over the pressure and flux spaces,

$$\begin{aligned} \|v_h\|_{X_h}^2 &= [v_h, v_h]_{X_h}, \\ \|q_h\|_{Q_h}^2 &= [q_h, q_h]_{Q_h}. \end{aligned}$$

The fact that  $\|\cdot\|_{X_h}$  is a norm follows from

$$s_*|E| \sum_{e \in \partial E} |v_E^e|^2 \leq [v_E, v_E]_E \leq S^*|E| \sum_{e \in \partial E} |v_E^e|^2 \quad \forall E \in \mathcal{T}_h, \forall v_E \in X_h(E), \quad (2.19)$$

which is a direct consequence of (2.11) in condition (S1). Note that (2.4) and (2.19) imply that for any  $v \in (H^1(E))^d$ ,

$$\|v^I\|_{X_{h,E}} \leq C\|v\|_{(H^1(E))^d}. \quad (2.20)$$

It is easy to see that (2.11) and (2.12) in (S1) imply the continuity of the bilinear form,

$$[u_h, v_h]_{X_h} \leq \frac{S^*}{s_*} \|u_h\|_{X_h} \|v_h\|_{X_h}. \quad (2.21)$$

We also define discrete  $H(\Omega; \text{div})$  norm

$$\|v_h\|_{\text{div}}^2 = \|v_h\|_{X_h}^2 + \|\mathcal{DIV} v_h\|_{Q_h}^2.$$

We clearly have that  $\forall q_h \in Q_h$  and  $\forall v_h \in X_h$ ,

$$[q_h, \mathcal{DIV} v_h]_{Q_h} \leq \|q_h\|_{Q_h} \|v_h\|_{\text{div}}.$$

Let  $Z_h$  represent the discrete divergence-free subspace of  $X_h$ ,

$$Z_h = \{v_h \in X_h \mid [\mathcal{DIV} v_h, q_h]_{Q_h} = 0, \forall q_h \in Q_h\}.$$

Note that since  $\mathcal{DIV} v_h \in Q_h$ ,  $v_h \in Z_h$  implies that  $\mathcal{DIV} v_h = 0$ . This immediately implies the coercivity of the flux bilinear operator,

$$[v_h, v_h]_{X_h} = \|v_h\|_{\text{div}}^2 \quad \forall v_h \in Z_h. \quad (2.22)$$

The inf-sup condition for the bilinear operator  $[\cdot, \mathcal{DIV} \cdot]_{Q_h}$ , has been shown in [37].

**Theorem 1** (inf-sup). *There exists a positive constant  $\beta$  independent of  $h$  such that, for any  $q_h \in Q_h$ ,*

$$\sup_{\{v_h \in X_h, v_h \neq 0\}} \frac{[\mathcal{DIV} v_h, q_h]_{Q_h}}{\|v_h\|_{\text{div}}} \geq \beta \|q_h\|_{Q_h}. \quad (2.23)$$

The existence and uniqueness of the solution  $(v_h, p_h)$  to (2.17)–(2.18) follows from (2.22) and Theorem 1, using the general theory for saddle-point problems [12].

## 2.2 Velocity Convergence

In this section we establish a first-order error estimate for the velocity variable. We will utilize an approximation result from [8]: for any element  $E$  and  $\phi \in H^2(E)$ , there exists a linear function  $\phi_E^1$  such that

$$\|\phi - \phi_E^1\|_{H^k(E)} \leq Ch_E^{m-k} |\phi|_{H^m(E)}, \quad k = 0, 1, \quad m = 1, 2. \quad (2.24)$$

We will also utilize the trace inequalities [2]:

$$\forall \chi \in H^1(E), \|\chi\|_{L^2(e)}^2 \leq C \left( h_E^{-1} \|\chi\|_{L^2(E)}^2 + h_E |\chi|_{H^1(E)}^2 \right), \quad (2.25)$$

$$\forall v \in (H^1(E))^d, \|v \cdot n_E^e\|_{L^2(e)}^2 \leq C \left( h_E^{-1} \|v\|_{(L^2(E))^d}^2 + h_E |v|_{(H^1(E))^d}^2 \right). \quad (2.26)$$

A combination of (2.24) and (2.25) implies that for any  $\phi \in H^2(E)$ , there exists a linear function  $\phi_E^1$  such that

$$\|\phi - \phi_E^1\|_{L^2(e)}^2 \leq Ch_E^3 |\phi|_{H^2(E)}^2. \quad (2.27)$$

The main result of this section is the following theorem bounding the velocity error.

**Theorem 2** (Velocity Estimate). *For the exact solution  $(u, p)$  of (2.1) and the MFD approximation  $(u_h, p_h)$  solving (2.17)–(2.18), assuming that  $p \in H^2(\Omega)$ , and  $u \in (H^1(\Omega))^d$ , there exists a positive constant  $C$ , independent of  $h$ , such that*

$$\|u^I - u_h\|_{X_h} \leq Ch(|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}). \quad (2.28)$$

*Proof.* Set  $v_h = u^I - u_h$ . Note that, using Lemma 1,

$$\mathcal{DIV} v_h = \mathcal{DIV} (u^I - u_h) = f^I - f^I = 0.$$

Thus,

$$\begin{aligned} \|u^I - u_h\|_{X_h}^2 &= [u^I - u_h, v_h]_{X_h} \\ &= [u^I, v_h]_{X_h} - [p_h, \mathcal{DIV} v_h]_{Q_h} \\ &= [u^I, v_h]_{X_h}. \end{aligned} \quad (2.29)$$

Let  $p^1$  be a piecewise linear function such that  $p^1|_E = p_E^1$ . By adding and subtracting  $(\bar{K}\nabla p^1)^I$ ,

$$[u^I, v_h]_{X_h} = \underbrace{[u^I + (\bar{K}\nabla p^1)^I, v_h]_{X_h}}_{I_1} - \underbrace{[(\bar{K}\nabla p^1)^I, v_h]_{X_h}}_{I_2}. \quad (2.30)$$

We now bound  $I_1$  and  $I_2$ . Starting with  $I_1$ ,

$$\begin{aligned}
|I_1| &\leq C \|(u + \bar{K} \nabla p^1)^I\|_{X_h} \|v_h\|_{X_h} \quad (\text{using (2.21)}) \\
&\leq C \left( \sum_{E \in \Omega} \sum_{e \in \partial E} \left( ((u + \bar{K}_E \nabla p^1)^I)_E^e \right)^2 |E| \right)^{1/2} \|v_h\|_{X_h} \quad (\text{using (2.19)}) \\
&= C \left( \sum_{E \in \Omega} \sum_{e \in \partial E} \left( \frac{1}{|e|} \int_e (u + \bar{K}_E \nabla p_E^1) \cdot n_E^e \, dS \right)^2 |E| \right)^{1/2} \|v_h\|_{X_h} \quad (\text{using (2.4)}) \\
&\leq C \left( \sum_{E \in \Omega} \sum_{e \in \partial E} \left( \frac{|E|}{|e|} \|(u + \bar{K}_E \nabla p_E^1) \cdot n_E^e\|_{L^2(e)}^2 \right)^{1/2} \right) \|v_h\|_{X_h} \\
&\leq C \left( \sum_{E \in \Omega} \sum_{e \in \partial E} \left( \|(u + \bar{K}_E \nabla p_E^1) \cdot n_E^e\|_{L^2(e)}^2 h_E \right)^{1/2} \right) \|v_h\|_{X_h} \quad (\text{using (2.2)}) \\
&\leq C \left( \sum_{E \in \Omega} \left( h_E^{-1} \|(u + \bar{K}_E \nabla p_E^1)\|_{(L^2(E))^d}^2 + h_E |u|_{(H^1(E))^d}^2 \right) h_E \right)^{1/2} \|v_h\|_{X_h} \\
&\quad (\text{using (2.26)}). \tag{2.31}
\end{aligned}$$

Taking the first term, we add and subtract  $K \nabla p_E^1$ ,

$$\begin{aligned}
\|u + \bar{K}_E \nabla p_E^1\|_{(L^2(E))^d} &\leq \|K \nabla (p - p_E^1)\|_{(L^2(E))^d} + \|(K - \bar{K}_E) \nabla p_E^1\|_{(L^2(E))^d} \\
&\leq C (h_E |p|_{H^2(E)} + h_E \|\nabla p_E^1\|_{(L^2(E))^d}) \quad (\text{using (2.6) and (2.24)}) \\
&\leq C (h_E |p|_{H^2(E)} + h_E |p|_{H^1(E)}), \tag{2.32}
\end{aligned}$$

where the last inequality follow from, using (2.24),

$$\|\nabla p_E^1\|_{(L^2(E))^d} \leq \|\nabla p\|_{(L^2(E))^d} + \|\nabla (p - p_E^1)\|_{(L^2(E))^d} \leq C |p|_{H^1(E)}.$$

Combining (2.31) and (2.32), we obtain

$$|I_1| \leq Ch (|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}) \|v_h\|_{X_h}.$$

Since  $\mathcal{DIV} v_h = 0$ , using condition  $(\widetilde{S2})$ , the expression for  $I_2$  becomes

$$I_2 = - \sum_{E \in \Omega} \sum_{e \in \partial E} v_E^e \int_e w_e p_E^1 \, dS. \tag{2.33}$$

Due to the continuity of  $p$ , we can subtract it from each element face in the summation above,

obtaining

$$\begin{aligned}
|I_2| &= \left| \sum_{E \in \Omega} \sum_{e \in \partial E} v_E^e \int_e w_e (p_E^1 - p) dS \right| \\
&\leq \sum_{E \in \Omega} \sum_{e \in \partial E} |e|^{1/2} |w_e^{\max}| |v_E^e| \|p_E^1 - p\|_{L^2(e)} \\
&\leq C \sum_{E \in \Omega} \left( |E| \sum_{e \in \partial E} |v_E^e|^2 \right)^{1/2} h_E |p|_{H^2(E)} \quad (\text{using (2.2) and (2.27)}) \\
&\leq Ch |p|_{H^2(\Omega)} \|v_h\|_{X_h} \quad (\text{using (S1)}). \tag{2.34}
\end{aligned}$$

Combining the bounds on  $I_1$  and  $I_2$  with (2.29) and (2.30) gives the desired estimate.  $\square$

### 2.2.1 Weighted Projection Operator

We also study the operator  $u^{Iw} \in X_h$  such that,

$$(u^{Iw})_E^e = \frac{1}{|e|} \int_e w_e u \cdot n_E^e dS.$$

Note that this weighted operator appears in the consistency condition ( $\widetilde{S2}$ ) and therefore it can be expected that the numerical flux  $u_h$  provides an improved approximation to  $u^{Iw}$ , compared to  $u^I$ . We study both errors in the numerical results. We have the following result.

**Corollary 1.** *Under the assumptions of Theorem 2,*

$$\|u^{Iw} - u_h\|_{X_h} \leq Ch(|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + \|u\|_{(H^1(\Omega))^d}).$$

*Proof.* The triangle inequality implies

$$\|u^{Iw} - u_h\|_{X_h} \leq \|u^{Iw} - u^I\|_{X_h} + \|u^I - u_h\|_{X_h}.$$

The second term is bounded in Theorem 2. Letting,

$$\bar{u} \cdot n_E^e = \int_e u \cdot n_E^e dS,$$

for the first term we have,

$$\begin{aligned}
\|u^{Iw} - u^I\|_{X_h}^2 &\leq S^* \sum_E |E| \sum_{e \in \partial E} |(u^{Iw} - u^I)_E^e|^2 \\
&= S^* \sum_E \sum_{e \in \partial E} \frac{|E|}{|e|^2} \left( \int_e w_e (u - \bar{u}) \cdot n_E^e \right)^2 \\
&\leq S^* (w_e^{\max})^2 \sum_E \sum_{e \in \partial E} \frac{|E|}{|e|} \int_e ((u - \bar{u}) \cdot n_E^e)^2 \\
&\leq Ch^2 \|u\|_{(H^1(\Omega))^d}^2 \quad (\text{using arguments similar to Lemma 3.14 in [49]}).
\end{aligned}$$

Combining the above bounds implies the statement of the corollary.  $\square$

### 2.3 Pressure Convergence

We start by defining a weighted operator  $I_w$  such that

$$(p^{I_w})_E = \frac{1}{|E|} \int_E w_E p \, dV.$$

Note that the operator is exact for constant functions. We first establish a pressure estimate with respect to the weighted operator  $(\cdot)^{I_w}$ . We then establish a bound for the regular projection operator  $(\cdot)^I$ .

**Theorem 3.** *Let  $(u, p)$  be the exact solution to (2.1), and let  $(u_h, p_h)$  be the MFD method's numerical approximation (2.17)-(2.18). Assuming that  $p \in H^2(\Omega)$ , and  $u \in (H^1(\Omega))^d$  there exists a constant  $C$  independent of  $h$ , such that*

$$\|p^{I_w} - p_h\|_{Q_h} \leq Ch(|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}). \quad (2.35)$$

*Proof.* From Theorem 1 we know that

$$\|p^{I_w} - p_h\|_{Q_h} \leq \frac{1}{\beta} \sup_{v_h \in X_h, v_h \neq 0} \frac{[\mathcal{D}\mathcal{I}\mathcal{V} v_h, p^{I_w} - p_h]_{Q_h}}{\|v_h\|_{\text{div}}}.$$

Adding and subtracting  $(p^1)^{I_w}$ , we get

$$[\mathcal{D}\mathcal{I}\mathcal{V} v_h, p^{I_w} - p_h]_{Q_h} = [\mathcal{D}\mathcal{I}\mathcal{V} v_h, (p - p^1)^{I_w}]_{Q_h} - [\mathcal{D}\mathcal{I}\mathcal{V} v_h, p_h]_{Q_h} + [\mathcal{D}\mathcal{I}\mathcal{V} v_h, (p^1)^{I_w}]_{Q_h}.$$

Using (2.17) for the second term and condition  $(\widetilde{S}2)$  for the third term, we have

$$\begin{aligned} [\mathcal{D}\mathcal{I}\mathcal{V} v_h, p^{I_w} - p_h]_{Q_h} &= \underbrace{[\mathcal{D}\mathcal{I}\mathcal{V} v_h, (p - p^1)^{I_w}]_{Q_h}}_{I_3} - \underbrace{[u_h, v_h]_{X_h}}_{I_4} \\ &\quad + \underbrace{\sum_E \sum_e u_E^e \int_e w_e p_E^1 \, dS}_{I_5} - \underbrace{\sum_E [(\bar{K}_E \nabla p_E^1)^I, v_h]_E}_{I_6}. \end{aligned}$$

We can bound  $I_3$  by

$$\begin{aligned} [\mathcal{D}\mathcal{I}\mathcal{V} v_h, (p - p^1)^{I_w}]_{Q_h} &= \sum_E (\mathcal{D}\mathcal{I}\mathcal{V} v_h, w_E (p - p_E^1))_{L^2(E)} \\ &\leq \sum_E \|\mathcal{D}\mathcal{I}\mathcal{V} v_h\|_{L^2(E)} \|w_E (p - p_E^1)\|_{L^2(E)} \\ &\leq \sum_E \|\mathcal{D}\mathcal{I}\mathcal{V} v_h\|_{L^2(E)} \|w_E\|_{L^\infty(E)} \|p - p_E^1\|_{L^2(E)} \\ &\leq w_E^{\max} \left( \sum_E \|\mathcal{D}\mathcal{I}\mathcal{V} v_h\|_{L^2(E)}^2 \right)^{1/2} \left( \sum_E \|p - p_E^1\|_{L^2(E)}^2 \right)^{1/2} \\ &\leq Ch |p|_{H^1(\Omega)} \|\mathcal{D}\mathcal{I}\mathcal{V} v_h\|_{Q_h} \quad (\text{using (2.24)}). \end{aligned}$$

Expression  $I_5$  is identical to  $I_2$  in (2.33), implying

$$|I_5| \leq Ch|p|_{H^2(\Omega)} \|v_h\|_{X_h}.$$

Taking  $I_4$  and  $I_6$ , and adding and subtracting  $u^I$ , we have

$$\begin{aligned} I_4 + I_6 &= [(\bar{K}\nabla p^1)^I + u^I, v_h]_{X_h} - [u^I - u_h, v_h]_{X_h}, \\ &= \tilde{I}_4 + \tilde{I}_6. \end{aligned}$$

Expression  $\tilde{I}_4$  is identical to  $I_1$  in (2.30), giving the bound

$$|\tilde{I}_4| \leq Ch(|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}) \|v_h\|_{X_h}.$$

Expression  $\tilde{I}_6$  is bounded by the velocity estimate in Theorem 2,

$$\begin{aligned} |\tilde{I}_6| &\leq C \|u^I - u_h\|_{X_h} \|v_h\|_{X_h} \\ &\leq Ch(|p|_{H^2(\Omega)} + |u|_{(H^1(\Omega))^d}) \|v_h\|_{X_h}. \end{aligned}$$

Combining the bounds on  $I_3 - I_6$  gives the desired result.  $\square$

We can now derive a bound based on the  $L^2$  projection operator  $(\cdot)^I$  defined in (2.3).

**Corollary 2.** *Under the assumptions of Theorem 3,*

$$\|p^I - p_h\|_{Q_h} \leq Ch(|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}). \quad (2.36)$$

*Proof.* We start with

$$\|p^I - p_h\|_{Q_h} \leq \|p^I - p^{I_w}\|_{Q_h} + \|p^{I_w} - p_h\|_{Q_h}.$$

The second term is bounded by Theorem 3. For the first term, we use the Bramble-Hilbert Lemma [8] and that the operator  $(\cdot)^{I_w}$  is exact for constants. Letting  $\bar{p}_E = \frac{1}{|E|} \int_E p \, dV$ , we have

$$\begin{aligned} \|p^I - p^{I_w}\|_{Q_h} &= \left( \sum_E \frac{1}{|E|} \left( \int_E w_E (\bar{p} - p) \, dV \right)^2 \right)^{1/2} \\ &\leq \left( \sum_E \frac{1}{|E|} \left( \int_E |w_E (\bar{p} - p)| \, dV \right)^2 \right)^{1/2} \\ &\leq \left( \sum_E \frac{1}{|E|} \left( \|w_E\|_{L^\infty(E)} \int_E |(\bar{p} - p)| \, dV \right)^2 \right)^{1/2} \\ &\leq w_E^{\max} \|\bar{p} - p\|_{L^2(\Omega)} \\ &\leq Ch|p|_{H^1(\Omega)}. \end{aligned}$$

Combining the above bounds implies the statement of the corollary.  $\square$

## 2.4 Superconvergence of Pressure

In this section we prove second-order convergence for  $\|p_h - p^I\|_{Q_h}$ . Note that  $\|p_h - p^I\|_{Q_h}$  is not in general superconvergent, i.e., there may not be second-order convergence at the centroids of the elements. This is confirmed by numerical experiments (Section 5).

We require the existence of two lifting operators  $R_E, \tilde{R}_E : X_h(E) \rightarrow H(E; \text{div})$  defined locally over each element and satisfying the following properties:

$$R_E(v_0^I) = v_0, \text{ for all constant vectors } v_0, \quad (2.37)$$

$$\nabla \cdot R_E(v_E) = \mathcal{D}\mathcal{T}\mathcal{V} v_E \text{ in } E, \quad (2.38)$$

$$\nabla \cdot \tilde{R}_E(v_E) = w_E \mathcal{D}\mathcal{T}\mathcal{V} v_E \text{ in } E, \quad (2.39)$$

$$\tilde{R}_E(v_E) \cdot n_E^e = w_e v_E^e \text{ on } e \subset \partial E, \quad (2.40)$$

$$\|R_E(v_E)\|_{(L^2(E))^d} \leq C \|v_E\|_{X_{h,E}}, \quad (2.41)$$

$$\|\tilde{R}_E(v_E)\|_{(L^2(E))^d} \leq C \|v_E\|_{X_{h,E}}. \quad (2.42)$$

**Lemma 2.** *The lifting operators  $R_E$  and  $\tilde{R}_E$  define a bilinear form,*

$$[u_E, v_E]_E = \int_E \bar{K}_E^{-1} R_E(u_E) \cdot \tilde{R}_E(v_E) \quad (2.43)$$

that satisfies condition  $(\widetilde{S2})$ .

*Proof.* We can demonstrate the result by observing that for all linear  $q^1$  and all  $v_E \in X_h(E)$ ,

$$\begin{aligned} [(\bar{K}_E \nabla q^1)^I, v_E]_E &= \int_E \bar{K}_E^{-1} R_E((\bar{K}_E \nabla q^1)^I) \cdot \tilde{R}_E(v_E) dV \\ &= \int_E \bar{K}_E^{-1} \bar{K}_E \nabla q^1 \cdot \tilde{R}_E(v_E) dV \quad (\text{using (2.37)}) \\ &= \sum_{e \in \partial E} \int_e q^1 \tilde{R}_E(v_E) \cdot n_E^e dS - \int_E q^1 \nabla \cdot \tilde{R}_E(v_E) dV \quad (\text{integration by parts}) \\ &= \sum_{e \in \partial E} v_E^e \int_e w_e q^1 dS - \int_E w_E q^1 \mathcal{D}\mathcal{T}\mathcal{V} v_E dV \quad (\text{using (2.40) and (2.39)}). \end{aligned} \quad (2.44)$$

□

We will need the following result proved in Lemma 3.3 in [37]:

**Lemma 3.** *For any  $v \in (H^1(E))^d$ , let  $v_0$  be a constant vector representing the  $L^2$  projection of  $v$  on  $E$ . Then, there exists a constant  $C$  independent of  $h$  such that*

$$\|v^I - v_0^I\|_{X_{h,E}} \leq Ch_E |v|_{(H^1(E))^d}. \quad (2.45)$$

The lifting operator  $\tilde{R}_E$  will not in general map constants to constants (as in (2.37)). However, we associate  $\tilde{R}_E$  with a tensor  $T_E$  with the property that for a constant vector  $v_0$ ,

$$\tilde{R}_E(v_0^I) = T_E v_0. \quad (2.46)$$

We require that the operator  $T_E$  is bounded:

$$\|T_E v\|_{(L^2(E))^d} \leq C \|v\|_{(L^2(E))^d}. \quad (2.47)$$

Note that (2.40) implies that for a constant vector  $v_0$ ,

$$T_E v_0 \cdot n_E^e = w_e v_0 \cdot n_E^e \text{ on } e \subset \partial E,$$

which, combined with (2.14) implies that

$$\|(T_E v_0 - v_0) \cdot n_E^e\|_{L^2(e)} \leq Ch \|v_0 \cdot n_E^e\|_{L^2(e)},$$

where  $C$  depends on  $|w_e|_{H^1, \infty(e)}$ , which is independent of  $h$ , due to (2.14)–(2.15). Since  $T_E$  can be defined to be in a finite dimensional space, the above inequality and a scaling argument imply that

$$\|T_E v_0 - v_0\|_{(L^2(E))^d} \leq Ch \|v_0\|_{(L^2(E))^d}. \quad (2.48)$$

We have the following results.

**Lemma 4.** For all  $v \in (H^1(E))^d$ ,

$$\|R_E(v^I) - v\|_{(L^2(E))^d} \leq Ch_E |v|_{(H^1(E))^d}, \quad (2.49)$$

$$\|\tilde{R}_E(v^I) - v\|_{(L^2(E))^d} \leq Ch_E \|v\|_{(H^1(E))^d}. \quad (2.50)$$

*Proof.* Let  $v_0$  be the  $L^2(E)$  projection of  $v$ . Using (2.37), (2.41) and (2.45), we have

$$\begin{aligned} \|R_E(v^I) - v\|_{(L^2(E))^d} &\leq \|R_E(v^I - v_0^I)\|_{(L^2(E))^d} + \|v_0 - v\|_{(L^2(E))^d} \\ &\leq Ch_E |v|_{(H^1(E))^d}, \end{aligned}$$

where we also used that

$$\|v - v_0\|_{(L^2(E))^d} \leq Ch_E |v|_{(H^1(E))^d}. \quad (2.51)$$

Similarly, using (2.42), (2.45), (2.46), (2.48) and (2.51),

$$\begin{aligned} \|\tilde{R}_E(v^I) - v\|_{(L^2(E))^d} &\leq \|\tilde{R}_E(v^I - v_0^I)\|_{(L^2(E))^d} + \|T_E v_0 - v_0\|_{(L^2(E))^d} + \|v_0 - v\|_{(L^2(E))^d} \\ &\leq Ch_E \|v\|_{(H^1(E))^d}. \end{aligned}$$

□

**Theorem 4.** Assuming the existence of lifting operators  $R_E$  and  $\tilde{R}_E$  with properties (2.37)–(2.42) and the choice of velocity bilinear form (2.43), then the the solution  $p_h$  to (2.17)–(2.18) satisfies

$$\|p_h - p^{Iw}\|_{Q_h} \leq Ch^2 (|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + \|u\|_{(H^1(\Omega))^d} + |f|_{H^1(\Omega)}). \quad (2.52)$$

*Proof.* We start by defining  $\varphi$  and  $\psi$  such that,

$$\varphi = K\nabla\psi \quad \text{in } \Omega, \quad (2.53)$$

$$\nabla \cdot \varphi = p_h - p^{Iw} \quad \text{in } \Omega, \quad (2.54)$$

$$\psi = 0 \quad \text{on } \partial\Omega. \quad (2.55)$$

In the above, by abuse of notation,  $p_h - p^{Iw}$  is identified with a piecewise constant function. Let  $\varphi_h = \varphi^I$ , and note that by Lemma 1,  $\varphi_h$  satisfies

$$\mathcal{DTV} \varphi_h = p_h - p^{Iw}.$$

We require  $H^2$ -regularity of problem (2.55). Conditions can be found in [24], which in this case can be satisfied by assuming convexity of  $\Omega$  and using (2.5). As a result we have,

$$\|\psi\|_{H^2(\Omega)} \leq C\|p_h - p^{Iw}\|_{L^2(\Omega)} = C\|p_h - p^{Iw}\|_{Q_h}. \quad (2.56)$$

We have

$$\begin{aligned} \|p_h - p^{Iw}\|_{Q_h}^2 &= [p_h - p^{Iw}, \mathcal{DTV} \varphi_h]_{Q_h} \\ &= [u_h, \varphi_h]_{X_h} - \int_{\Omega} p \nabla \cdot \tilde{R}(\varphi_h) \, dV \quad (\text{using (2.17) and (2.39)}) \\ &= [u_h, \varphi_h]_{X_h} + \int_{\Omega} \nabla p \cdot \tilde{R}(\varphi_h) \, dV \quad (\text{integration by parts}) \\ &= \sum_E \left[ \int_E \bar{K}_E^{-1} (R_E(u_E) - u) \cdot \tilde{R}_E(\varphi_E) \, dV + \int_E (\bar{K}_E^{-1} - K^{-1}) u \cdot \tilde{R}_E(\varphi_E) \, dV \right] \\ &= \sum_E \left[ \int_E (\bar{K}_E^{-1} - K^{-1}) (R_E(u_E) - u) \cdot \tilde{R}_E(\varphi_E) \, dV \right. \\ &\quad \left. + \int_E K^{-1} (R_E(u_E) - u) \cdot \tilde{R}_E(\varphi_E) \, dV \right. \\ &\quad \left. + \int_E (\bar{K}_E^{-1} - K^{-1}) u \cdot \tilde{R}_E(\varphi_E) \, dV \right] \\ &= J_1 + J_2 + J_3 \end{aligned}$$

For  $J_1$ , using (2.7), we have,

$$\begin{aligned} |J_1| &\leq Ch \|R(u_h) - u\|_{(L^2(\Omega))^d} \|\tilde{R}(\varphi)\|_{(L^2(\Omega))^d} \\ &\leq Ch^2 (|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}) \|p_h - p^{Iw}\|_{Q_h}, \end{aligned}$$

where for the second step, we used

$$\begin{aligned} \|R(u_h) - u\|_{(L^2(E))^d} &\leq C \|R(u_h - u^I)\|_{(L^2(\Omega))^d} + \|R(u^I) - u\|_{(L^2(\Omega))^d} \\ &\leq C \|u_h - u^I\|_{X_h} + \|R(u^I) - u\|_{(L^2(\Omega))^d} \quad (\text{using (2.41)}) \\ &\leq Ch (|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}) \quad (\text{by Theorem 2 and (2.49)}), \end{aligned} \quad (2.57)$$

as well as

$$\begin{aligned}
\|\tilde{R}(\varphi_h)\|_{(L^2(\Omega))^d} &\leq C\|\varphi_h\|_{X_h} \text{ (using (2.42))} \\
&\leq C\|\varphi\|_{(H^1(\Omega))^d} \text{ (using (2.20))} \\
&\leq C\|\psi\|_{H^2(\Omega)} \text{ (using (2.5))} \\
&\leq C\|p_h - p^{Iw}\|_{Q_h} \text{ (using (2.56))}.
\end{aligned} \tag{2.58}$$

For  $J_2$ , we start by adding and subtracting  $\varphi$ ,

$$\begin{aligned}
J_2 &= \int_{\Omega} K^{-1}(R(u_h) - u) \cdot (\tilde{R}(\varphi_h) - \varphi) dV + \int_{\Omega} K^{-1}(R(u_h) - u) \cdot \varphi dV \\
&= J_{21} + J_{22}.
\end{aligned}$$

For  $J_{21}$ , using (2.50) and (2.57), we have

$$|J_{21}| \leq Ch^2(|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d})\|p_h - p^{Iw}\|_{Q_h}.$$

We can bound  $J_{22}$  by noting that,

$$\begin{aligned}
J_{22} &= \int_{\Omega} K^{-1}(R(u_h) - u) \cdot K\nabla\psi dV \text{ (using (2.53))} \\
&= \int_{\Omega} (R(u_h) - u) \cdot \nabla\psi dV \\
&= - \int_{\Omega} \nabla \cdot (R(u_h) - u)\psi dV \\
&= \int_{\Omega} (f^I - f)(\psi - \psi^I) dV.
\end{aligned}$$

From here we have by the Bramble-Hilbert Lemma,

$$|J_{22}| \leq Ch^2|f|_{H^1(\Omega)}\|p_h - p^{Iw}\|_{Q_h}$$

For  $J_3$ , we add and subtract  $\varphi$ , obtaining

$$\begin{aligned}
J_3 &= \sum_E \int_E \left[ (\bar{K}_E^{-1} - K^{-1})u \cdot (\tilde{R}_E(\varphi_E) - \varphi) dV + \int_E (\bar{K}_E^{-1} - K^{-1})u \cdot \varphi dV \right] \\
&= J_{31} + J_{32}.
\end{aligned}$$

Using (2.7) and (2.50) and an argument similar to (2.58),  $J_{31}$  is bounded by,

$$|J_{31}| \leq Ch^2\|u\|_{(L^2(\Omega))^d}\|p_h - p^{Iw}\|_{Q_h}.$$

For expression  $J_{32}$ , we have,

$$\begin{aligned}
J_{32} &= \sum_E \int_E (\bar{K}_E^{-1} - K^{-1})u \cdot K\nabla\psi dV \\
&= \sum_E \int_E (K - \bar{K}_E)\bar{K}_E^{-1}u \cdot K^{-1}K\nabla\psi dV \\
&= \sum_E \int_E (K - \bar{K}_E)\bar{K}_E^{-1}(u - u_0) \cdot \nabla\psi dV + \int_E (K - \bar{K}_E)\bar{K}_E^{-1}u_0 \cdot (\nabla\psi - (\nabla\psi)_0) dV
\end{aligned}$$

An application of (2.6) and the Bramble-Hilbert Lemma gives

$$|J_{32}| \leq Ch^2 \|u\|_{(H^1(\Omega))^d} \|p_h - p^{Iw}\|_{Q_h}$$

Combining all the expressions gives the desired result.  $\square$

#### 2.4.1 Quadrature

An alternative approach to pressure superconvergence is to use an approximate quadrature rule for the velocity bilinear form. In this case, we will restrict  $[\cdot, \cdot]_E$  to symmetric form. Define  $\sigma_E(K^{-1}; u_E, v_E)$  as

$$\sigma_E(K^{-1}; u_E, v_E) = [u_E, v_E]_E - \int_E K^{-1} R_E(u_E) \cdot \tilde{R}_E(v_E) dV. \quad (2.59)$$

**Theorem 5.** *Assume the existence of lifting operators  $R_E$  and  $\tilde{R}_E$  with properties (2.37)–(2.42) and that the choice of a symmetric velocity bilinear form  $[\cdot, \cdot]_E$  satisfies for all  $u, v \in (H^1(E))^d$*

$$|\sigma_E(K^{-1}, (u^I)_E, (v^I)_E)| \leq Ch_E^2 \|u\|_{(H^1(E))^d} \|v\|_{(H^1(E))^d}. \quad (2.60)$$

Then the solution  $p_h$  to problem (2.17)–(2.18) satisfies

$$\|p_h - p^{Iw}\|_{Q_h} \leq Ch^2 (|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + \|u\|_{(H^1(\Omega))^d} + |f|_{H^1(\Omega)}).$$

*Proof.* We utilize again the auxiliary problem (2.53)–(2.55) and recall that  $\varphi_h = \varphi^I$ . We have

$$\begin{aligned} \|p_h - p^{Iw}\|_{Q_h}^2 &= [p_h - p^{Iw}, \mathcal{DIV} \varphi_h]_{Q_h} \\ &= [u_h, \varphi_h]_{X_h} - \int_{\Omega} p \nabla \cdot \tilde{R}(\varphi_h) dV \quad (\text{using (2.17) and (2.39)}) \\ &= [u_h, \varphi_h]_{X_h} + \int_{\Omega} \nabla p \cdot \tilde{R}(\varphi_h) dV \quad (\text{integration by parts}) \\ &= [u_h, \varphi_h]_{X_h} - \int_{\Omega} K^{-1} u \cdot \tilde{R}(\varphi_h) dV \pm \int_{\Omega} K^{-1} R(u^I) \cdot \tilde{R}(\varphi_h) dV \\ &= [u_h - u^I, \varphi_h]_{X_h} + \sum_E \sigma_E(K^{-1}; (u^I)_E, (\varphi^I)_E) + \int_{\Omega} K^{-1} (R(u^I) - u) \cdot \tilde{R}(\varphi_h) dV \\ &= J_4 + J_5 + J_6. \end{aligned}$$

For  $J_4$  we have

$$\begin{aligned} J_4 &= [u_h - u^I, \varphi_h + (\bar{K} \nabla \varphi^1)^I]_{X_h} - [u_h - u^I, (\bar{K} \nabla \varphi^1)^I]_{X_h} \\ &= J_{41} + J_{42}. \end{aligned}$$

For  $J_{41}$ , using the same method as for  $I_1$  in Theorem 2, we have

$$\begin{aligned} |J_{41}| &\leq Ch \|u_h - u^I\|_{X_h} \|\psi\|_{H^2(\Omega)} \\ &\leq Ch \|u_h - u^I\|_{X_h} \|p^{Iw} - p_h\|_{Q_h}. \end{aligned}$$

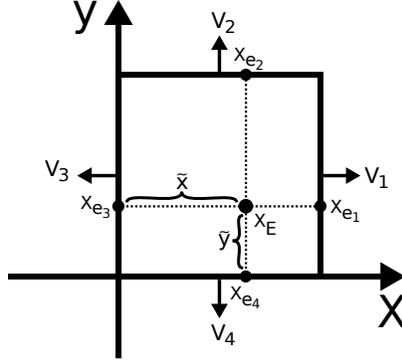


Figure 1: The reference element for the lifting operator  $\tilde{R}$ .

Since the bilinear form is symmetric, we can directly apply the argument for  $I_2$  from Theorem 2 to term  $J_{42}$ :

$$\begin{aligned} |J_{42}| &\leq Ch \|u_h - u^I\|_{X_h} |\psi|_{H^2(\Omega)} \\ &\leq Ch \|u_h - u^I\|_{X_h} \|p^{I_w} - p_h\|_{Q_h}. \end{aligned}$$

Combining  $J_{41}$  and  $J_{42}$ , and using the velocity estimate from Theorem 2, we obtain

$$|J_4| \leq Ch^2 (|p|_{H^2(\Omega)} + |p|_{H^1(\Omega)} + |u|_{(H^1(\Omega))^d}) \|p^{I_w} - p_h\|_{Q_h}.$$

For  $J_5$ , assumption (2.59) gives

$$\begin{aligned} |J_5| &\leq Ch^2 \|u\|_{(H^1(\Omega))^d} \|\varphi\|_{(H^1(\Omega))^d} \\ &\leq Ch^2 \|u\|_{(H^1(\Omega))^d} \|p^{I_w} - p_h\|_{Q_h}. \end{aligned}$$

Term  $J_6$  is bounded similarly to term  $J_2$  in Theorem 4:

$$|J_6| \leq Ch^2 (|u|_{(H^1(\Omega))^d} + |f|_{H^1(\Omega)}) \|p^{I_w} - p_h\|_{Q_h}.$$

Combining all the terms gives the desired result.  $\square$

#### 2.4.2 Lifting Operators for a Square Element

We present an example of lifting operators  $R_E$  and  $\tilde{R}_E$  satisfying (2.37)-(2.42) over a square reference element  $E = [0, 1]^2$ , see Figure 1, where  $x_E = (\tilde{x}, \tilde{y})^T$ ,  $x_{e_2} = x_{e_4} = \tilde{x}$ ,  $x_{e_1} = x_{e_3} = \tilde{y}$ . For operator  $R_E$ , we use the standard lowest order Raviart-Thomas ( $RT_0$ ) interpolant:

$$R_E(v_E) = \begin{pmatrix} v_3 + (v_1 - v_3)x \\ v_4 + (v_2 - v_4)y \end{pmatrix},$$

where  $v_i = v_E^{e_i}$ . It is well known [12] that the  $RT_0$  interpolant satisfies (2.37), (2.38) and (2.41). The lifting operator  $\tilde{R}_E$  is defined as

$$\tilde{R}_E(v_E) = \begin{pmatrix} (v_3 + (v_1 - v_3)W_x(x))w_y(y) \\ (v_4 + (v_2 - v_4)W_y(y))w_x(x) \end{pmatrix},$$

with

$$W_x(0) = 0, \quad W_x(1) = 1, \quad \frac{\partial}{\partial x} W_x = w_x, \quad (2.61)$$

$$W_y(0) = 0, \quad W_y(1) = 1, \quad \frac{\partial}{\partial y} W_y = w_y, \quad (2.62)$$

where the functions  $w_x(x)$  and  $w_y(y)$  are the weighting functions on the horizontal and vertical faces, respectively. The corresponding volume weighting function  $w_E$  is

$$w_E(x, y) = w_x(x)w_y(y).$$

In this form, the lifting operator  $\tilde{R}_E$  satisfies conditions (2.39) and (2.40):

$$\tilde{R}_E(v_E) \cdot n_1|_{x=1} = v_1 w_y,$$

$$\tilde{R}_E(v_E) \cdot n_3|_{x=0} = v_3 w_y,$$

$$\tilde{R}_E(v_E) \cdot n_2|_{y=1} = v_2 w_x,$$

$$\tilde{R}_E(v_E) \cdot n_4|_{y=0} = v_4 w_x,$$

and

$$\begin{aligned} \nabla \cdot \tilde{R}_E(v_E) &= w_x w_y (v_1 - v_3) + w_x w_y (v_2 - v_4) \\ &= w_E \mathcal{D}\mathcal{I}\mathcal{V} v_E. \end{aligned}$$

We now are left with defining functions  $w_y$  and  $w_x$ , as well as  $W_x$  and  $W_y$ . We construct a function of the form

$$W_x(x) = \frac{a}{2}x^2 + bx,$$

and  $w_x$  is of the form

$$w_x(x) = ax + b.$$

Functions  $w_x$  and  $w_y$  can be constructed directly by satisfying condition (2.14), (2.15), that is:

$$\int_0^1 w_x \, dS = 1, \quad (2.63)$$

$$\int_0^1 x w_x \, dS = \tilde{x}. \quad (2.64)$$

Note that this choice satisfies conditions (2.61) for  $W_x$ , since

$$1 = \int_0^1 w_x dS = \frac{a}{2} + b = W_x(1).$$

Equations (2.63)–(2.64) result in the linear system

$$\begin{pmatrix} 1/2 & 1 \\ 1/3 & 1/2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ \tilde{x} \end{pmatrix},$$

which has a solution

$$a = -6 + 12\tilde{x}, \quad b = 4 - 6\tilde{x}.$$

The form for  $w_y$  is similar. The resulting function  $w_E$  satisfies,

$$\int_E w_E dV = 1$$

and

$$\int_E \begin{pmatrix} x \\ y \end{pmatrix} w_E = \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = x_E.$$

It is easy to check that (2.42) holds and that (2.46) holds with

$$T_E = \begin{pmatrix} w_y & 0 \\ 0 & w_x \end{pmatrix}.$$

Therefore the lifting operators satisfy all conditions needed for the pressure superconvergence in Theorem 4 with velocity inner product defined by (2.43).

For the alternative approach to pressure superconvergence in Theorem 5, consider the case of diagonal permeability with

$$\bar{K}_E = \begin{pmatrix} k_x & 0 \\ 0 & k_y \end{pmatrix}.$$

We can define the velocity inner product via (2.10) with a diagonal quadrature matrix  $\mathbf{M}_E$ :

$$\mathbf{M}_E^{(1,1)} = \frac{1 - \tilde{x}}{k_x}, \quad \mathbf{M}_E^{(2,2)} = \frac{1 - \tilde{y}}{k_y}, \quad \mathbf{M}_E^{(3,3)} = \frac{\tilde{x}}{k_x}, \quad \mathbf{M}_E^{(4,4)} = \frac{\tilde{y}}{k_y}. \quad (2.65)$$

Note that the positivity of the entries ensures satisfaction of condition (S1) and that condition (S2) is satisfied by the above choice of inner product.

Notice that for constant  $u_0$  we have

$$\sigma_E(\bar{K}_E^{-1}; u_0, v_E) = 0, \quad (2.66)$$

which can be observed by setting  $u_0 = (1, 0)^T$  and  $v_E = (1, 0, 0, 0)$ :

$$\begin{aligned} \int_E \bar{K}_E^{-1} R_E(u_0) \cdot \tilde{R}_E(v_E) dV &= \int_E \bar{K}_E^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} W_x w_y \\ 0 \end{pmatrix} dV \\ &= \frac{1 - \tilde{x}}{k_x} = [u_0, v_E]_E. \end{aligned}$$

The same argument can be extended to all other combinations of  $u_0$  and  $v_E$ . Due to non-symmetry of  $\sigma_E$ , in general

$$\sigma_E(\bar{K}_E^{-1}; v_E, u_0) \neq 0.$$

However, we can proceed by first defining an operator  $\hat{T}_E$ ,

$$\hat{T}_E = \begin{pmatrix} \hat{w}_x & 0 \\ 0 & \hat{w}_y \end{pmatrix},$$

with

$$\begin{aligned} \int_e x \hat{w}_x dS &= 1 - \tilde{x}, \\ \int_e \hat{w}_x dS &= 1, \end{aligned}$$

and similarly for  $\hat{w}_y$ . We now observe that for any  $v_E$  and  $u_0$ ,

$$\begin{aligned} [v_E, u_0]_E &= \int_E \bar{K}_E^{-1} R_E(v_E) \cdot \hat{T}_E \tilde{R}_E(u_0) \\ &= \int_E \bar{K}_E^{-1} R_E(v_E) \cdot \hat{T}_E T_E u_0, \end{aligned} \tag{2.67}$$

since for  $v_E = (1, 0, 0, 0)$  and  $u_0 = (1, 0)$  we have

$$\begin{aligned} \int_E \bar{K}_E^{-1} R_E(v_E) \cdot \hat{T}_E T_E u_0 &= \int_E k_x^{-1} x \hat{w}_x w_y \\ &= \frac{1 - \tilde{x}}{k_x} = [v_E, u_0]_E, \end{aligned}$$

and for  $v_E = (0, 0, 1, 0)$  and  $u_0 = (1, 0)$ ,

$$\begin{aligned} \int_E \bar{K}_E^{-1} R_E(v_E) \cdot \hat{T}_E T_E u_0 &= \int_E k_x^{-1} (1 - x) \hat{w}_x w_y \\ &= \frac{\tilde{x}}{k_x} = [v_E, u_0]_E, \end{aligned}$$

and similarly for all other combinations of  $v_E$  and  $u_0$ .

We note that on any  $E \in \mathcal{T}_h$ , since

$$\int_E \hat{w}_x = \int_E \hat{w}_y = |E|,$$

and using (2.47), we have that for all constant vectors  $v_0$ ,

$$\|\hat{T}_E T_E v_0 - T_E v_0\|_{(L^2(E))^2} \leq Ch_E \|v_0\|_{(L^2(E))^2}. \tag{2.68}$$

**Lemma 5.** *The choice of velocity inner product (2.65) satisfies condition (2.60).*

*Proof.* Given  $u \in (H^1(E))^d$  and  $v \in (H^1(E))^d$ , let  $u_E = u_E^I$  and  $v_E = v_E^I$ . We have

$$\begin{aligned}\sigma_E(K^{-1}; u_E, v_E) &= \sigma_E(\bar{K}_E^{-1}; u_E, v_E) + \int_E (\bar{K}_E^{-1} - K^{-1}) R_E(u_E) \cdot \tilde{R}_E(v_E) dV \\ &= \tilde{J}_1 + \tilde{J}_2.\end{aligned}$$

For  $\tilde{J}_1$  we have, using (2.66) and (2.67),

$$\begin{aligned}\tilde{J}_1 &= [u_E, v_E]_E - \int_E \bar{K}_E^{-1} R_E(u_E) \cdot \tilde{R}_E(v_E) dV, \\ &= [u_E - u_0, v_E - v_0]_E - \int_E \bar{K}_E^{-1} R_E(u_E - u_0) \cdot (\tilde{R}_E(v_E) - \hat{T}_E T_E v_0) dV \\ &= [u_E - u_0, v_E - v_0]_E - \int_E \bar{K}_E^{-1} R_E(u_E - u_0) \cdot \tilde{R}_E(v_E - v_0) dV \\ &\quad - \int_E \bar{K}_E^{-1} R_E(u_E - u_0) \cdot (T_E v_0 - \hat{T}_E T_E v_0) dV.\end{aligned}$$

Using Lemma 3, Lemma 4, (2.68) and (2.48), we obtain

$$|\tilde{J}_1| \leq Ch_E^2 \|u\|_{(H^1(E))^2} \|v\|_{(H^1(E))^2}.$$

Expression  $\tilde{J}_2$  can be divided into

$$\begin{aligned}\tilde{J}_2 &= \int_E (\bar{K}_E^{-1} - K^{-1}) R_E(u_E - u_0) \cdot \tilde{R}_E(v_E) dV \\ &\quad + \int_E (\bar{K}_E^{-1} - K^{-1}) R_E(u_0) \cdot \tilde{R}_E(v_E - v_0) dV \\ &\quad + \int_E (\bar{K}_E^{-1} - K^{-1}) R_E(u_0) \cdot \tilde{R}_E(v_0) dV \\ &= \tilde{J}_{21} + \tilde{J}_{22} + \tilde{J}_{23}\end{aligned}\tag{2.69}$$

We can bound  $\tilde{J}_{21} + \tilde{J}_{22}$  using (2.7), (2.41)-(2.42) and (2.51),

$$|\tilde{J}_{21} + \tilde{J}_{22}| \leq Ch_E^2 \|v\|_{(H^1(E))^2} \|u\|_{(H^1(E))^2}.\tag{2.70}$$

For expression  $\tilde{J}_{23}$ , we have, setting  $\overline{K^{-1}T_E} = \frac{1}{|E|} \int_E K^{-1} T_E$  and using (2.6),

$$\begin{aligned}|\tilde{J}_{23}| &= \left| \int_E (\bar{K}_E - K) \bar{K}_E^{-1} u_0 \cdot K^{-1} T_E v_0 dV \right| \\ &= \left| \int_E (\bar{K}_E - K) \bar{K}_E^{-1} u_0 \cdot (K^{-1} T_E - \overline{K^{-1}T_E}) v_0 dV \right| \\ &\leq Ch_E^2 \|u\|_{(L^2(E))^2} \|v\|_{(L^2(E))^2}.\end{aligned}\tag{2.71}$$

Combining the above expressions gives the desired result.  $\square$

### 3 Matrix Construction

We have explicitly defined the  $\mathcal{DIV}$  operator, as well as the pressure inner product  $[\cdot, \cdot]_{Q_h}$ . What remains is to define, for the general case, the velocity bilinear  $[\cdot, \cdot]_{X_h}$ . Recall that this bilinear is defined relative to the local bilinears,

$$[u_h, v_h]_{X_h} = \sum_{E \in \mathcal{T}_h} [u_h, v_h]_E.$$

The discretization yields a convergent solution if the inner product satisfies the stability (S1) and consistency ( $\widetilde{S2}$ ) conditions. The second condition is the main focus of constructing the appropriate inner product.

Define matrix  $\mathbf{R}_E \in \mathbb{R}^{k_E \times d}$  as

$$\mathbf{R}_E = \begin{pmatrix} |e_1|(x_{e_1} - x_E)^T \\ \vdots \\ |e_{k_E}|(x_{e_{k_E}} - x_E)^T \end{pmatrix}, \quad (3.1)$$

and matrix  $\mathbf{N}_E \in \mathbb{R}^{k_E \times d}$  as

$$\mathbf{N}_E = \begin{pmatrix} (\bar{K}_E n_{e_1})^T \\ \vdots \\ (\bar{K}_E n_{e_{k_E}})^T \end{pmatrix}, \quad (3.2)$$

see Figure 2. Condition ( $\widetilde{S2}$ ) can be satisfied with  $\mathbf{M}_E$  defined to satisfy

$$\mathbf{M}_E \mathbf{N}_E = \mathbf{R}_E. \quad (3.3)$$

Note that by setting  $w_E = 1$  and  $w_e = 1$  we retrieve the original definition of the MFD method [14], in which case,  $x_E$  corresponds to the centroid of element  $E$ , and  $x_e$  would correspond to the centroid of face  $e$ . Also note that we do not require explicit construction of the weighting functions in order to build the linear system.

The authors of [14] demonstrate how to construct  $\mathbf{M}_E$  to satisfy (3.3). This is done by first defining  $\mathbf{C}_E \in \mathbb{R}^{k_E \times k_E - d}$  such that

$$\mathbf{N}_E^T \mathbf{C}_E = 0.$$

In [14], they note that without weights,

$$\mathbf{N}_E^T \mathbf{R}_E = |E| \bar{K}_E \quad (3.4)$$

so they define  $\mathbf{M}_0$  as

$$\mathbf{M}_0 = \frac{1}{|E|} \mathbf{R}_E \bar{K}_E^{-1} \mathbf{R}_E^T.$$

The result is that, for any positive-definite  $\mathbf{U}_E \in \mathbb{R}^{(K_E-d) \times (K_E-d)}$ ,

$$\mathbf{M}_E = \mathbf{M}_0 + \mathbf{C}_E \mathbf{U}_E \mathbf{C}_E^T, \quad (3.5)$$

satisfies (3.3). Due to the modification ( $\widetilde{S2}$ ) of condition (S2), equality (3.4) no longer holds. However, we can still proceed by following the same form proposed by [34],

$$\mathbf{M}_0 = \mathbf{R}_E (\mathbf{R}_E^T \mathbf{N}_E)^{-1} \mathbf{R}_E^T.$$

Doing so, we see that (3.5) implies (3.3), since

$$\mathbf{M}_0 \mathbf{N}_E = \mathbf{R}_E (\mathbf{R}_E^T \mathbf{N}_E)^{-1} \mathbf{R}_E^T \mathbf{N}_E = \mathbf{R}_E.$$

Note that  $\mathbf{R}_E^T \mathbf{N}_E$  automatically reduces to  $|E| \bar{K}_E^{-1}$  in the case when no boundary points are shifted. In general,  $\mathbf{R}_E^T \mathbf{N}_E$  may not always be invertible, in which case we can use an alternative form for  $\mathbf{M}_0$  found in [37],

$$\mathbf{M}_0 = \mathbf{R}_E (\mathbf{N}_E^T \mathbf{N}_E)^{-1} \mathbf{N}_E^T.$$

## 4 Finite Volume Methods and the MFD Method

Finite Volume Methods (FVM) are based on an application of the divergence theorem to each element  $E$  in the domain,

$$\int_E \nabla \cdot u \, dV = \sum_{e \in \partial E} \int_e u \cdot n_e^e \, dS \approx \frac{1}{|E|} \sum_{e \in \partial E} |e| u_E^e.$$

Finite volume discretizations are distinguished by how the flux ( $u_E^e$ ) is approximated. In the case of porous media applications, the flux of the fluid is a function of the fluid pressure. Each element in the domain has a single, piecewise constant pressure degree of freedom ( $p_E$ ), and each face in the mesh has a single flux representing the normal component of the velocity across that face ( $u_E^e$ ). The flux is then related to the pressure via a function

$$u_E^e = F(p_h).$$

The choice of function  $F$  is what distinguishes a particular FV method. When the velocity at face  $e$  is approximated using *two* adjacent pressures, we have a two-point flux approximation (TPFA). When more than two pressures are used, the method is referred to as a multi-point flux approximation (MPFA).

The MFD method, much like the related MFE method, produces a saddle-point problem with both velocity and pressure unknowns:

$$\begin{pmatrix} M & -\mathcal{DIV}^* \\ \mathcal{DIV} & 0 \end{pmatrix} \begin{pmatrix} u_h \\ p_h \end{pmatrix} = \begin{pmatrix} 0 \\ f^I \end{pmatrix}.$$

One can observe the relation between FVM and the MFD method by explicitly expressing the velocity unknown as a function of pressure:

$$u_h = -M^{-1} \mathcal{DIV}^* p_h = F_{\text{MFD}}(p_h).$$

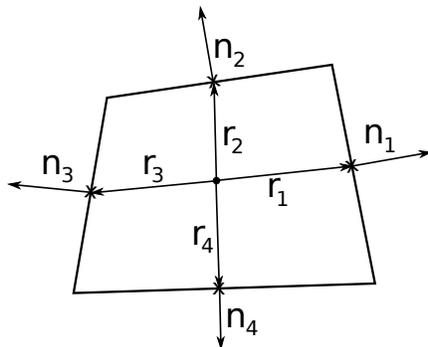


Figure 2: The vectors  $r$  are constructed by connecting  $x_E$  with the appropriate point  $x_e$ . These vectors make up the rows of matrix  $\mathbf{R}_E$ ,  $(|e_1 r_1^T, \dots |e_4 r_4^T)$ . The rows of matrix  $\mathbf{N}_E$  are  $((\bar{K}_E n_1)^T, \dots (\bar{K}_E n_4)^T)$ .

The nature of the relation between  $p_h$  and  $u_h$  is a direct consequence of the structure of  $M^{-1}$ . In general, the matrix  $M^{-1}$  is dense, causing the velocity at every face to be a function of all the pressure unknowns. However, constructing a diagonal matrix  $M$  leads trivially to a diagonal  $M^{-1}$ . In this case, due to the structure of matrix  $D\mathcal{I}\mathcal{V}^*$ , velocities are a linear function of two pressures, resulting in a TPFA scheme.

There is a simple geometric criterion that indicates when diagonality can be achieved in the MFD method. Recall that matrix  $M$  is a global matrix formed from the summation of local matrices,  $\mathbf{M}_E$ ,

$$M = \sum_E \mathbf{M}_E.$$

Due to the consistency condition ( $\widetilde{S2}$ ),  $\mathbf{M}_E$  must satisfy the relation

$$\mathbf{M}_E \mathbf{N}_E = \mathbf{R}_E. \quad (4.1)$$

Recall from (3.1) that  $\mathbf{R}_E$  is a  $k_E \times d$  matrix with rows corresponding to the vectors  $|e|(x_e - x_E)^T$ . From (3.2), the rows of matrix  $\mathbf{N}_E$  correspond to the  $K$ -normal components to face  $e$ , see Figure 2. When the rows of  $\mathbf{N}_E$  are collinear to the rows of  $\mathbf{R}_E$ , a simple scaling of the rows of  $\mathbf{N}_E$  satisfies (4.1), *i.e.*, a diagonal matrix  $\mathbf{M}_E$  suffices. Therefore, our objective is to construct matrices  $\mathbf{R}_E$  with rows collinear to the rows of  $\mathbf{N}_E$ . The original MFD method sets the point  $x_E$  to the centroid of the cells and the points  $x_e$  to the centroids of the faces. Our generalization allows the point  $x_E$  to be shifted on the interior of the cell, and the points  $x_e$  to be shifted on the plane of the faces. By shifting these points, we can establish collinearity of the rows of  $\mathbf{R}_E$  and  $\mathbf{N}_E$ , resulting in a diagonal matrix  $\mathbf{M}_E$ .

Next, we show that this can be achieved for general Voronoi meshes when  $K$  is a scalar function. A Voronoi diagram is a tessellation of  $\mathbb{R}^d$  relative to a set of points known as generators.

**Definition 1.** *Voronoi Diagram.* Given a set of generating points,  $V = \{V_i \in \mathbb{R}^d\}$ , we define

the Voronoi tessellation  $\mathcal{T}_h = \{E_i\}$  as

$$E_i = \{x \in \mathbb{R}^d \mid d(x, V_i) \leq d(x, V_j) \forall j \neq i\}.$$

The set of generating points uniquely defines a Voronoi diagram. Note that, in this definition, the domain is infinite. For our purposes we will focus on what is known as a *bounded Voronoi diagram*, which is defined over a bounded domain  $\Omega$  as follows.

**Definition 2.** *Bounded Voronoi Diagram.* Given a domain  $\Omega \subset \mathbb{R}^d$ , and a set of generating points  $V = \{V_i \in \mathbb{R}^d \mid V_i \in \Omega\}$ , we define the bounded Voronoi tessellation  $\mathcal{T}_h = \{E_i\}$  of  $\Omega$  by

$$E_i = \{x \in \Omega \mid d(x, V_i) \leq d(x, V_j) \forall j \neq i\}.$$

Since we are only concerned with bounded domains, we refer to a *bounded Voronoi diagram* simply as a *Voronoi diagram*. Each cell  $E_i$  is called a Voronoi polygon/polyhedron. We say that two Voronoi cells (and their generating points) are adjacent if they share a non-trivial Voronoi face. It is a direct consequence of the definition of Voronoi diagrams that the line joining two adjacent generating points is always perpendicular to the common face between them and bisected by the plane of the face. For two adjacent points,  $V_i$  and  $V_j$ , with associated face  $e$ , we refer to the midpoint between them as  $b_e$ .

Bounded diagrams introduce certain challenges because fundamental properties of the Voronoi diagram are broken at the boundary. For example, a non-convex boundary could lead to non-convex cells. An extreme case can be considered by selecting a single generating point in an arbitrarily shaped domain. In order to simplify the problem, we only consider convex boundaries. In addition, we require that the orthogonality of the diagram is maintained at the domain boundary.

**Lemma 6.** For a Voronoi polyhedron with isotropic, piecewise constant permeability ( $\bar{K}_E = \kappa_E I$ ), by setting  $x_E = V_E$  and  $x_e = b_e$ , a diagonal matrix  $\mathbf{M}_E$  can be constructed that satisfies both conditions (S1) and ( $\widetilde{S2}$ ).

*Proof.* Let us denote the rows of  $\mathbf{R}_E$  and  $\mathbf{N}_E$  by  $R_e^T$  and  $N_e^T$ , respectively. Since the line joining two adjacent generating points is orthogonal to the Voronoi face between them, the vectors  $R_e$  are collinear to  $N_e$ . That is,

$$R_e = \frac{|e| \|R_e\|}{\kappa_E} N_e,$$

where  $\|\cdot\|$  is the Euclidean norm on  $\mathbb{R}^d$ . Therefore, a diagonal matrix  $\mathbf{M}_E$  with choice of entries  $(\mathbf{M}_E)_{ii} = \frac{|e_i| \|R_{e_i}\|}{\kappa_E}$  satisfies condition ( $\widetilde{S2}$ ).  $\square$

Note that the point  $b_e$  may fall outside of the boundary faces, see Figure 3. This, however, presents no problems for our definition, as we allow the function  $w_e$  to be negative, which can shift the boundary point  $x_e$  outside of the face.

In porous media applications, it is common to use the so-called 2.5-dimensional Voronoi mesh. These diagrams are constructed by forming a two-dimensional Voronoi diagram. Cells are then vertically extruded into three-dimensional prisms. A 2.5-dimensional Voronoi mesh allows for slightly greater flexibility in permeability tensor when establishing TPFA.

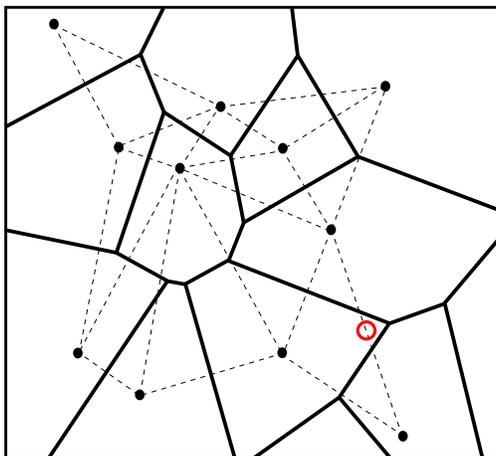


Figure 3: A two-dimensional Voronoi diagram (solid line) and the lines joining adjacent Voronoi generating points (dashed line), which corresponds to the Delaunay triangulation. The dashed lines are always perpendicular to and bisected by the corresponding edges of the Voronoi diagram. The red circle illustrates an example of how the intersection of the two lines may occur outside of the face boundaries.

**Lemma 7.** *Given a mesh that is a Voronoi diagram in the  $x$ - $y$  direction and orthogonally extruded in the  $z$  direction, let*

$$K_E = \begin{pmatrix} \kappa_x & 0 & 0 \\ 0 & \kappa_y & 0 \\ 0 & 0 & \kappa_z \end{pmatrix},$$

with  $\kappa_x = \kappa_y$ . Then there exists a diagonal matrix  $\mathbf{M}_E$  that satisfies (S1) and  $(\widetilde{S2})$ .

*Proof.* In this case, a diagonal  $M_E$  can be constructed by setting

$$(\mathbf{M}_E)_{ii} = \frac{|e_i| \|R_{e_i}\|}{\kappa_x}$$

for faces  $e$  facing in the  $x$ - $y$  direction, and

$$(\mathbf{M}_E)_{ii} = \frac{|e_i| \|R_{e_i}\|}{\kappa_z}$$

for faces on the  $z$ -plane. □

## 5 Numerical Results

In this section we present numerical results for our proposed generalization of the MFD method. The permeability tensor  $\bar{K}_E$  is computed at the centroid of each cell  $E$ . Given a set of points

$z_E \in E, \forall E \in \mathcal{T}_h$ , define

$$\text{PressErr}(z_E) = \left( \sum_E |E| (p(z_E) - p_E)^2 \right)^{1/2}.$$

Note that

$$\begin{aligned} \|p^{I_w} - p_h\|_{Q_h} &= \left( \sum_E |E| ((p^{I_w})_E - p_E)^2 \right)^{1/2} \\ &\approx \left( \sum_E |E| (p(x_E) - p_E)^2 \right)^{1/2} \\ &= \text{PressErr}(x_E). \end{aligned}$$

Note that for linear  $p|_E$ ,  $p_E^{I_w} = p(x_E)$ , therefore the above is a second-order approximation. Similarly,

$$\begin{aligned} \|p^I - p_h\|_{Q_h} &= \left( \sum_E |E| ((p^I)_E - p_E)^2 \right)^{1/2} \\ &\approx \left( \sum_E |E| (p(C_E) - p_E)^2 \right)^{1/2} \\ &= \text{PressErr}(C_E), \end{aligned}$$

which is also a second order approximation.

We can establish similar norms for the velocity errors. Given a set of points  $z_e \in e$ , define

$$\text{VelErr}(z_e) = \left( \sum_E |E| \sum_{e \in \partial E} (u(z_e) \cdot n_E^e - u_E^e)^2 \right)^{1/2}.$$

Note that

$$\begin{aligned} \|u^{I_w} - u_h\|_{X_h} &\approx \left( \sum_E |E| \sum_{e \in \partial E} ((u^{I_w})_E^e - u_E^e)^2 \right)^{1/2} \\ &\approx \left( \sum_E |E| \sum_{e \in \partial E} ((u(b_e) \cdot n_E^e - u_E^e)^2) \right)^{1/2} \\ &= \text{VelErr}(b_e). \end{aligned}$$

Note that this is a second order approximation, since for linear  $u \cdot n_E^e$  on  $e$ ,

$$(u^{I_w})_E^e = u(b_e) \cdot n_E^e.$$

Similarly,

$$\begin{aligned} \|u^I - u_h\|_{X_h} &\approx \left( \sum_E |E| \sum_{e \in \partial E} ((u^I)_E^e - u_E^e)^2 \right)^{1/2} \\ &\approx \left( \sum_E |E| \sum_{e \in \partial E} ((u(c_e) \cdot n_E^e - u_E^e)^2) \right)^{1/2} \\ &= \text{VelErr}(c_e). \end{aligned}$$

For each mesh, we test four different methods based on choices of points  $x_E$  and  $x_e$  in  $(\widetilde{S2})$ :

- Case 1:  $x_E = C_E, x_e = c_e$ ,
- Case 2:  $x_E = V_E, x_e = c_e$ ,
- Case 3:  $x_E = C_E, x_e = b_e$ ,
- Case 4:  $x_E = V_E, x_e = b_e$ .

These four cases are illustrated in Figure 4. In all cases we compute the pressure errors for  $z_E = C_E$ , the centroid of  $E$ , and  $z_E = V_E$ , the generating point for the Voronoi diagram and the velocity errors for  $z_e = c_e$ , the centroid of  $e$ , and  $z_e = b_e$ , the Voronoi bisection point.

For Case 1 and Case 3, we have

$$\begin{aligned} \text{PressErr}(C_E) &= \text{PressErr}(x_E) \approx \|p^{Iw} - p_h\|_{Q_h} = \|p^I - p_h\|_{Q_h}, \\ \text{PressErr}(V_E) &\text{ is not directly related to either error.} \end{aligned}$$

For Case 2 and Case 4, we have

$$\begin{aligned} \text{PressErr}(C_E) &\approx \|p^I - p_h\|_{Q_h}, \\ \text{PressErr}(V_E) &= \text{PressErr}(V_E) \approx \|p^{Iw} - p_h\|_{Q_h}. \end{aligned}$$

For Case 1 and Case 2, we have

$$\begin{aligned} \text{VelErr}(c_e) &= \text{VelErr}(x_e) \approx \|u^{Iw} - u_h\|_{X_h} = \|u^I - u_h\|_{X_h}, \\ \text{VelErr}(b_e) &\text{ is not directly related to either error.} \end{aligned}$$

For Case 3 and Case 4, we have

$$\begin{aligned} \text{VelErr}(c_e) &\approx \|u^I - u_h\|_{X_h}, \\ \text{VelErr}(b_e) &\approx \|u^{Iw} - u_h\|_{X_h}. \end{aligned}$$

The theory predicts for Case 1 and Case 3 the following pressure convergence rates:

$$\begin{aligned} \text{PressErr}(C_E) &= O(h^2), \\ \text{PressErr}(V_E) &= O(h). \end{aligned}$$

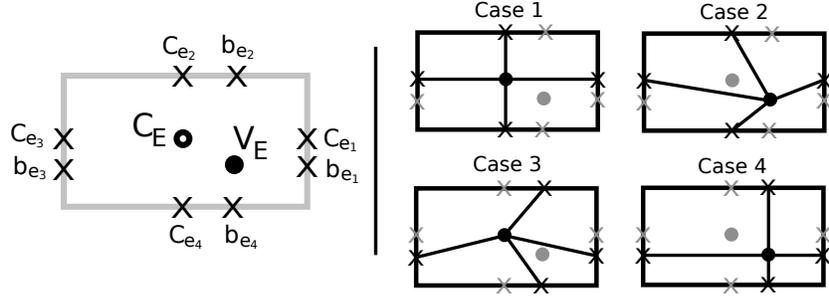


Figure 4: The four cases are based on different locations for placing points  $x_e$  and  $x_E$ . The point  $C_E$  represents the centroid of the cell  $E$ , and the point  $c_e$  represents the centroid of the face  $e$ . The point  $V_E$  represents the Voronoi generating point for cell  $E$ , and the point  $b_e$  represents the bisection point of the line joining two adjacent Voronoi generating points.

For Case 2 and Case 4, we expect the following convergence rates:

$$\begin{aligned} \text{PressErr}(C_E) &= O(h), \\ \text{PressErr}(V_E) &= O(h^2). \end{aligned}$$

We expect at least  $O(h)$  for all velocity errors. Recall from Section 2.2 that we expect better accuracy for  $\|u^{I_w} - u_h\|_{X_h}$  compared to  $\|u^I - u_h\|_{X_h}$ , which translates to better accuracy for  $\text{VelErr}(c_e)$  in Cases 1 and 2 and for  $\text{VelErr}(b_e)$  in Cases 3 and 4.

## 5.1 Two-Dimensional Problems

We present results confirming convergence of the modified MFD formulation for (2.1) over two-dimensional meshes. We consider the following manufactured solution from [14] over the unit square domain  $\Omega = [0, 1]^2$ ,

$$p(x, y, z) = x^3 y^2 + x \sin(2\pi xy) \sin(2\pi y) + 1,$$

and a full permeability tensor

$$K = \begin{pmatrix} (x+1)^2 + y^2 & -xy \\ -xy & (x+1)^2 \end{pmatrix}.$$

For all the example in this work, the matrix  $U_E$  is chosen to be

$$U_E = \frac{|E|}{\text{trace}(K_E)} \mathbb{I}.$$

### 5.1.1 Rectangular Grids

We carry out the results on the meshes in Figure 6, see also Figure 7 for an example of the computed pressure solution and the analytic solution. These meshes are generated by perturbing

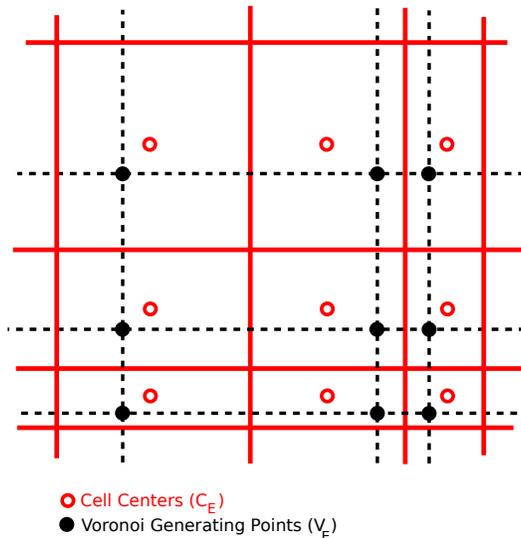


Figure 5: An illustration of the cell centered points and the Voronoi generating points used for the two-dimensional rectangular numerical experiments. In this case, the rectangular mesh is also a Voronoi diagram.

evenly spaced points with spacing  $h$  by

$$\xi_l = lh + \frac{3}{50} |\sin(4\pi lh)|, \quad \xi = x, y.$$

These points serve as the shifted location  $V_E$  for point  $x_E$ . The mesh is then constructed as a Voronoi diagram using these points as generating points (see Figure 5). This is equivalent to constructing a rectangular mesh in a point-centered fashion, where faces are placed midway between two adjacent points. The centroids of the cells formed by this procedure are the points  $C_E$ , see Figure 5.

As predicted, we observe  $O(h^2)$  convergence for the pressure at the points  $x_E$ , i.e., at points  $C_E$  for Cases 1 and 3, and at points  $V_E$  for Cases 2 and 4. We observe a deterioration in the convergence rate of approximately  $O(h^{3/2})$  at the points that are not used in the construction of matrix  $\mathbf{M}_E$ . With respect to the velocity error, we observe superconvergence of at least  $O(h^{3/2})$  in all cases. As expected, the convergence is somewhat better at the points  $x_e$ , i.e. at points  $c_e$  for Cases 1 and 2, and points  $b_e$  for Cases 3 and 4.

### 5.1.2 Voronoi Grids

We test the method using randomly generated Voronoi diagrams. The diagrams are constructed by selecting uniformly distributed random points in the domain as generating points. The mesh is refined by selecting a larger number of randomly generated points unrelated to the previous

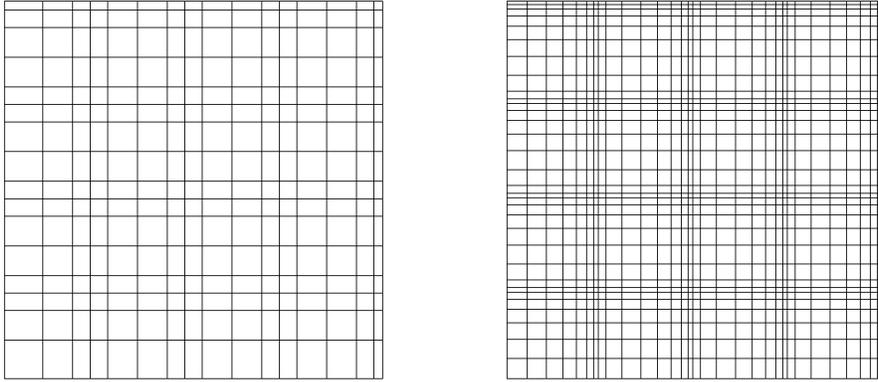


Figure 6: The rectangular meshes used in the convergence study.

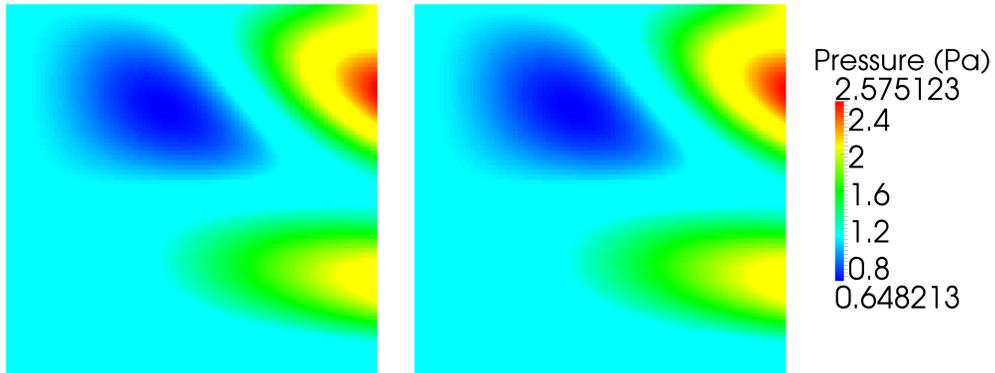


Figure 7: The MFD method pressure solution (left) and the analytic pressure solution (right) for using rectangular meshes.

| (Case 1: $x_E = C_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
|-----------------------------------|-------------------|-------|-------------------|-------|-----------------|-------|-----------------|-------|
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 1.129e-02         | —     | 2.729e-02         | —     | 6.270e-01       | —     | 7.277e-01       | —     |
| 32                                | 3.217e-03         | 1.81  | 8.850e-03         | 1.62  | 2.068e-01       | 1.60  | 2.413e-01       | 1.60  |
| 64                                | 8.616e-04         | 1.90  | 2.766e-03         | 1.68  | 6.191e-02       | 1.74  | 7.524e-02       | 1.68  |
| 128                               | 2.216e-04         | 1.96  | 8.903e-04         | 1.63  | 1.753e-02       | 1.82  | 2.297e-02       | 1.71  |
| 256                               | 5.660e-05         | 1.97  | 2.969e-04         | 1.58  | 4.829e-03       | 1.86  | 7.088e-03       | 1.70  |

| (Case 2: $x_E = V_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
|-----------------------------------|-------------------|-------|-------------------|-------|-----------------|-------|-----------------|-------|
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 2.439e-02         | —     | 1.385e-02         | —     | 9.220e-01       | —     | 1.020e+00       | —     |
| 32                                | 7.948e-03         | 1.61  | 3.638e-03         | 1.92  | 2.728e-01       | 1.76  | 3.028e-01       | 1.75  |
| 64                                | 2.558e-03         | 1.63  | 9.377e-04         | 1.96  | 7.654e-02       | 1.83  | 8.526e-02       | 1.83  |
| 128                               | 8.481e-04         | 1.59  | 2.374e-04         | 1.98  | 2.110e-02       | 1.86  | 2.398e-02       | 1.83  |
| 256                               | 2.893e-04         | 1.55  | 6.214e-05         | 1.93  | 5.849e-03       | 1.85  | 6.910e-03       | 1.80  |

| (Case 3: $x_E = C_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
|-----------------------------------|-------------------|-------|-------------------|-------|-----------------|-------|-----------------|-------|
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 1.358e-02         | —     | 2.899e-02         | —     | 9.206e-01       | —     | 8.245e-01       | —     |
| 32                                | 4.154e-03         | 1.71  | 9.510e-03         | 1.61  | 3.177e-01       | 1.54  | 2.754e-01       | 1.58  |
| 64                                | 1.134e-03         | 1.87  | 2.931e-03         | 1.70  | 9.395e-02       | 1.76  | 8.025e-02       | 1.78  |
| 128                               | 2.929e-04         | 1.95  | 9.253e-04         | 1.66  | 2.719e-02       | 1.79  | 2.258e-02       | 1.83  |
| 256                               | 7.493e-05         | 1.97  | 3.038e-04         | 1.60  | 8.007e-03       | 1.76  | 6.335e-03       | 1.83  |

| (Case 4: $x_E = V_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
|-----------------------------------|-------------------|-------|-------------------|-------|-----------------|-------|-----------------|-------|
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 2.352e-02         | —     | 1.272e-02         | —     | 8.712e-01       | —     | 8.048e-01       | —     |
| 32                                | 7.501e-03         | 1.65  | 3.184e-03         | 2.00  | 2.844e-01       | 1.62  | 2.436e-01       | 1.72  |
| 64                                | 2.438e-03         | 1.62  | 7.902e-04         | 2.01  | 8.755e-02       | 1.70  | 7.079e-02       | 1.78  |
| 128                               | 8.231e-04         | 1.57  | 1.965e-04         | 2.01  | 2.716e-02       | 1.69  | 2.070e-02       | 1.77  |
| 256                               | 2.842e-04         | 1.53  | 5.004e-05         | 1.97  | 8.694e-03       | 1.64  | 6.237e-03       | 1.73  |

Table 1: Convergence rates for the four test cases with rectangular grids in two dimensions. The results demonstrate at least first-order convergence across the board, and second-order convergence of the pressure variable when the error is calculated at the point  $x_E$ , as predicted by theory. We consistently find a slightly better velocity approximation when calculating the velocity error at point  $x_e$ .

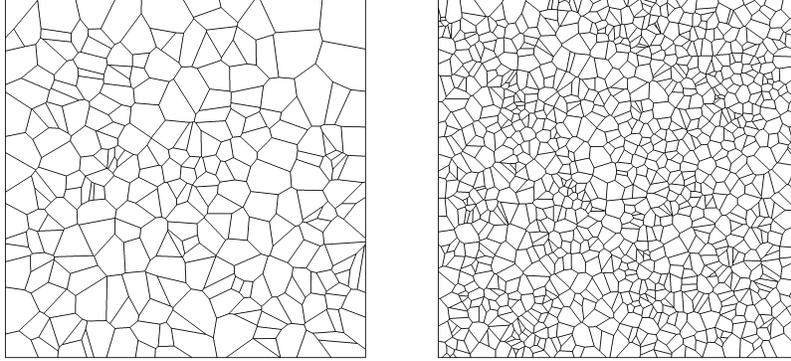


Figure 8: The random Voronoi meshes used in the convergence study.

mesh. Examples of the meshes used can be seen in Figure 8. An example of the computed pressure solutions along with the analytic solution can be seen in Figure 9.

We tested the four different cases, and the results can be seen in Table 2. We observe second-order convergence occurs at the location of point  $x_E$ . Shifting of the point on the interior changes where the solution is most accurate. As predicted by the theory, we consistently maintain at least first-order convergence of both pressure and flux independent of shifting  $x_E$  and  $x_e$ .

## 5.2 Three-Dimensional Problems

We now consider problems defined in three-dimensions. We solve the following manufactured problem over the unit cube domain  $\Omega = [0, 1]^3$ ,

$$p(x, y, z) = x^3 y^2 z + x \sin(2\pi xy) \sin(2\pi yz) \sin(2\pi z) + 1,$$

and a full permeability tensor

$$K = \begin{pmatrix} 1 + y^2 + z^2 & -xy & -xz \\ -xy & 1 + x^2 + z^2 & -yz \\ -xz & -yz & 1 + x^2 + y^2 \end{pmatrix}.$$

We follow the same four test cases outlined in the previous section. The matrix  $U_E$  is chosen to be

$$U_E = \frac{|E|}{\text{trace}(K_E)} \mathbb{I}.$$

| (Case 1: $x_E = C_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
|-----------------------------------|-------------------|-------|-------------------|-------|-----------------|-------|-----------------|-------|
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 1.470e-02         | —     | 3.615e-02         | —     | 2.004e+00       | —     | 2.435e+00       | —     |
| 32                                | 4.529e-03         | 1.70  | 1.835e-02         | 0.98  | 1.072e+00       | 0.90  | 1.181e+00       | 1.04  |
| 64                                | 1.150e-03         | 1.98  | 8.410e-03         | 1.13  | 4.963e-01       | 1.11  | 5.673e-01       | 1.06  |
| 128                               | 3.158e-04         | 1.86  | 4.239e-03         | 0.99  | 2.475e-01       | 1.00  | 2.933e-01       | 0.95  |
| 256                               | 7.641e-05         | 2.05  | 2.153e-03         | 0.98  | 1.221e-01       | 1.02  | 1.420e-01       | 1.05  |
| (Case 2: $x_E = V_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 4.339e-02         | —     | 2.712e-02         | —     | 2.290e+00       | —     | 2.717e+00       | —     |
| 32                                | 1.919e-02         | 1.18  | 6.545e-03         | 2.05  | 1.079e+00       | 1.09  | 1.167e+00       | 1.22  |
| 64                                | 8.541e-03         | 1.17  | 1.858e-03         | 1.82  | 4.834e-01       | 1.16  | 5.503e-01       | 1.08  |
| 128                               | 4.261e-03         | 1.00  | 5.388e-04         | 1.79  | 2.442e-01       | 0.99  | 2.845e-01       | 0.95  |
| 256                               | 2.156e-03         | 0.98  | 1.303e-04         | 2.05  | 1.168e-01       | 1.06  | 1.349e-01       | 1.08  |
| (Case 3: $x_E = C_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 2.219e-02         | —     | 4.208e-02         | —     | 2.735e+00       | —     | 2.378e+00       | —     |
| 32                                | 6.578e-03         | 1.75  | 1.927e-02         | 1.13  | 1.266e+00       | 1.11  | 1.099e+00       | 1.11  |
| 64                                | 1.547e-03         | 2.09  | 8.490e-03         | 1.18  | 5.992e-01       | 1.08  | 5.285e-01       | 1.06  |
| 128                               | 4.051e-04         | 1.93  | 4.250e-03         | 1.00  | 2.982e-01       | 1.01  | 2.649e-01       | 1.00  |
| 256                               | 9.297e-05         | 2.12  | 2.154e-03         | 0.98  | 1.480e-01       | 1.01  | 1.297e-01       | 1.03  |
| (Case 4: $x_E = V_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 16                                | 3.702e-02         | —     | 2.088e-02         | —     | 2.632e+00       | —     | 2.362e+00       | —     |
| 32                                | 1.899e-02         | 0.96  | 6.239e-03         | 1.74  | 1.263e+00       | 1.06  | 1.05e+00        | 1.163 |
| 64                                | 8.483e-03         | 1.16  | 1.653e-03         | 1.92  | 5.719e-01       | 1.14  | 4.855e-01       | 1.12  |
| 128                               | 4.244e-03         | 1.00  | 3.996e-04         | 2.05  | 2.884e-01       | 0.99  | 2.452e-01       | 0.99  |
| 256                               | 2.153e-03         | 0.98  | 9.858e-05         | 2.02  | 1.404e-01       | 1.04  | 1.177e-01       | 1.06  |

Table 2: Convergence rates for the four test cases with random Voronoi grids. We consistently observe first-order convergence for all cases, as predicted by the theory. We also observe second-order convergence of pressure when the error is calculated at  $x_E$ .

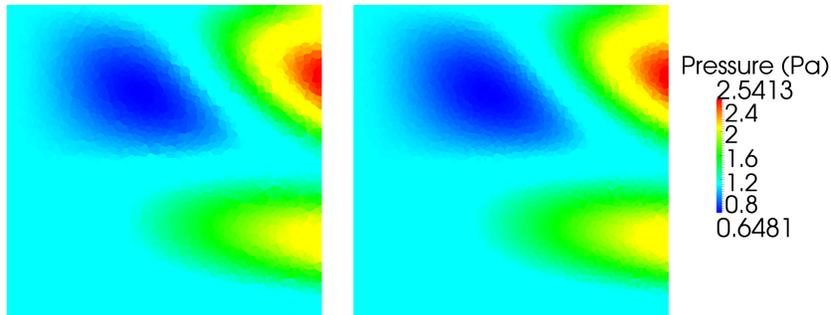


Figure 9: The MFD method pressure solution (left) and the analytic pressure solution computed at the cell centroids (right) using two-dimensional Voronoi diagrams.

### 5.2.1 Rectangular Grids

We carry out the tests on the meshes in Figure 10. These meshes are generated by perturbing points evenly spaced by  $h$  by,

$$\xi_l = lh + \frac{3}{50} |\sin(4\pi lh)|, \quad \xi = x, y, z.$$

The mesh is then constructed as a Voronoi diagram using these points as generating points. This is equivalent to constructing a rectangular mesh in a point-centered fashion, where faces are placed midway between two adjacent points. The results of these experiments are shown in Table 3. Examples of the pressure solution can be seen in Figure 11. Much like the two-dimensional problem, in all cases we have established at least first-order convergence, and often the method exhibits superconvergence. We observe again a second-order convergence of the pressure at the point  $x_E$ , as well as more accurate velocities at the points  $x_e$ .

### 5.2.2 Voronoi Grids

We test the method over three-dimensional Voronoi meshes. The meshes are generated by randomly selecting points uniformly over the domain. An example of the produced meshes can be seen in Figure 12. The pressure solution can be seen in Figure 13. The convergence results are shown in Table 4. Much like the two-dimensional case, we find that the pressure solution is most accurate at the points  $x_E$ .

## 6 Conclusions

The MFD method defines a family of discretizations on a very general set of polyhedral meshes. Solving over general polyhedra allows for better representation of complex geometric features.

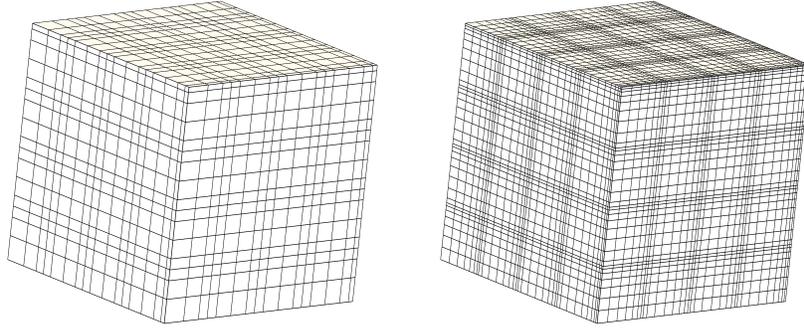


Figure 10: The rectangular meshes used in the convergence study. The two meshes correspond to  $h = \frac{1}{16}$  and  $\frac{1}{32}$ .

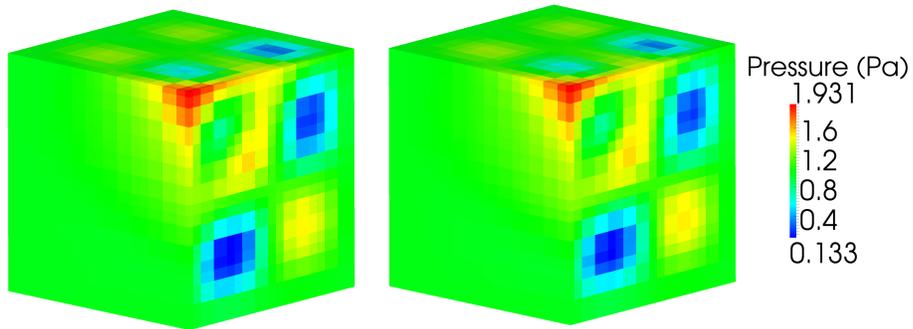


Figure 11: Comparison between the MFD method solution (left) and the exact solution evaluated at the cell centroids (right) over a three dimensional domain with full permeability tensor.

| (Case 1: $x_E = C_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
|-----------------------------------|-------------------|-------|-------------------|-------|-----------------|-------|-----------------|-------|
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 7.778e-02         | —     | 1.117e-01         | —     | 4.724e+00       | —     | 4.913e+00       | —     |
| 8                                 | 2.856e-02         | 1.45  | 5.370e-02         | 1.06  | 1.543e+00       | 1.61  | 1.640e+00       | 1.58  |
| 16                                | 1.059e-02         | 1.45  | 2.671e-02         | 1.01  | 6.376e-01       | 1.27  | 7.362e-01       | 1.16  |
| 32                                | 3.037e-03         | 1.79  | 8.717e-03         | 1.62  | 2.112e-01       | 1.59  | 2.446e-01       | 1.59  |
| 64                                | 8.184e-04         | 1.90  | 2.738e-03         | 1.67  | 6.340e-02       | 1.74  | 7.637e-02       | 1.68  |
| (Case 2: $x_E = V_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 8.610e-02         | —     | 9.978e-02         | —     | 4.648e+00       | —     | 4.831e+00       | —     |
| 8                                 | 3.663e-02         | 1.23  | 4.763e-02         | 1.07  | 2.799e+00       | 0.73  | 2.826e+00       | 0.77  |
| 16                                | 2.621e-02         | 0.48  | 1.681e-02         | 1.50  | 1.334e+00       | 1.07  | 1.405e+00       | 1.01  |
| 32                                | 8.694e-03         | 1.60  | 5.071e-03         | 1.72  | 4.350e-01       | 1.61  | 4.549e-01       | 1.63  |
| 64                                | 2.773e-03         | 1.65  | 1.421e-03         | 1.84  | 1.293e-01       | 1.75  | 1.347e-01       | 1.76  |
| (Case 3: $x_E = C_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 8.086e-02         | —     | 1.139e-01         | —     | 5.008e+00       | —     | 5.067e+00       | —     |
| 8                                 | 3.216e-02         | 1.33  | 5.537e-02         | 1.04  | 2.130e+00       | 1.23  | 2.100e+00       | 1.27  |
| 16                                | 1.751e-02         | 0.88  | 3.089e-02         | 0.84  | 1.285e+00       | 0.73  | 1.221e+00       | 0.78  |
| 32                                | 5.574e-03         | 1.65  | 1.018e-02         | 1.60  | 4.550e-01       | 1.50  | 4.262e-01       | 1.52  |
| 64                                | 1.572e-03         | 1.82  | 3.117e-03         | 1.71  | 1.387e-01       | 1.71  | 1.297e-01       | 1.72  |
| (Case 4: $x_E = V_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 7.944e-02         | —     | 9.446e-02         | —     | 4.350e+00       | —     | 4.357e+00       | —     |
| 8                                 | 3.164e-02         | 1.33  | 4.289e-02         | 1.14  | 2.250e+00       | 0.95  | 2.151e+00       | 1.02  |
| 16                                | 2.318e-02         | 0.45  | 1.186e-02         | 1.85  | 9.051e-01       | 1.31  | 8.400e-01       | 1.36  |
| 32                                | 7.442e-03         | 1.64  | 2.950e-03         | 2.01  | 2.938e-01       | 1.62  | 2.541e-01       | 1.73  |
| 64                                | 2.427e-03         | 1.62  | 7.279e-04         | 2.02  | 8.971e-02       | 1.71  | 7.335e-02       | 1.79  |

Table 3: Convergence rates for the four test cases with rectangular grids in three-dimensions. We consistently find at least first-order convergence for both pressure and velocity. Also note that the pressure solution is most accurate when the error is calculated using the  $x_E$  point. The velocity exhibits a slightly more accurate solution when the error is calculated at  $x_e$ .

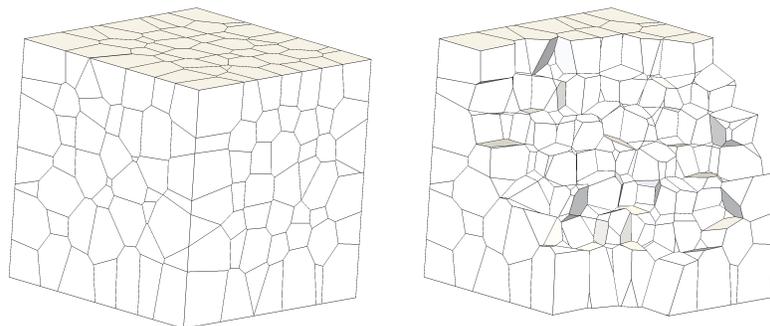


Figure 12: Unstructured three-dimensional Voronoi diagrams. The plot on the left is the full domain of the convergence test problem in which we can only see the outer faces of the boundary. The plot on the right is a slice of the mesh showing the inner structure of the diagram.

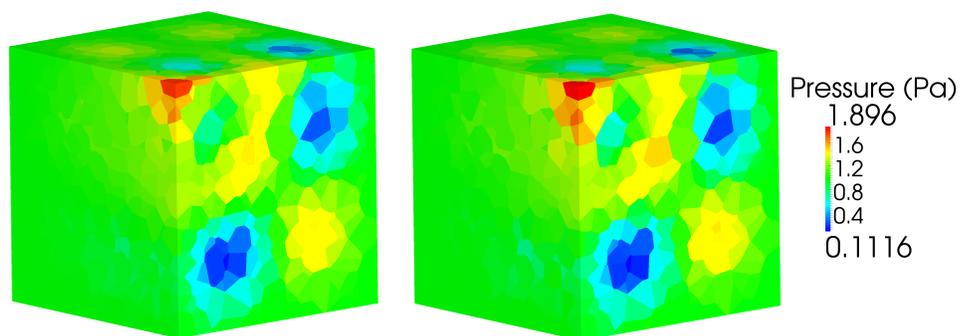


Figure 13: Pressure solution over a three-dimensional Voronoi diagram. The plot on the left represents the MFD method's approximation to pressure, and the plot on the right is the analytic solution computed at the cell centroids.

| (Case 1: $x_E = C_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
|-----------------------------------|-------------------|-------|-------------------|-------|-----------------|-------|-----------------|-------|
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 2.207e-01         | —     | 2.114e-01         | —     | 1.066e+01       | —     | 1.166e+01       | —     |
| 8                                 | 6.574e-02         | 1.75  | 8.806e-02         | 1.26  | 6.853e+00       | 0.64  | 6.635e+00       | 0.81  |
| 16                                | 2.493e-02         | 1.40  | 3.395e-02         | 1.37  | 3.579e+00       | 0.94  | 3.619e+00       | 0.87  |
| 32                                | 7.083e-03         | 1.82  | 1.304e-02         | 1.38  | 1.621e+00       | 1.14  | 1.643e+00       | 1.14  |
| (Case 2: $x_E = V_E, x_e = c_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 3.102e-01         | —     | 3.042e-01         | —     | 1.415e+01       | —     | 1.474e+01       | —     |
| 8                                 | 1.163e-01         | 1.42  | 1.085e-01         | 1.49  | 8.140e+00       | 0.80  | 8.022e+00       | 0.88  |
| 16                                | 4.172e-02         | 1.48  | 3.340e-02         | 1.70  | 3.775e+00       | 1.11  | 3.843e+00       | 1.06  |
| 32                                | 1.438e-02         | 1.54  | 9.108e-03         | 1.87  | 1.666e+00       | 1.18  | 1.724e+00       | 1.16  |
| (Case 3: $x_E = C_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 2.423e-01         | —     | 2.391e-01         | —     | 1.176e+01       | —     | 1.236e+01       | —     |
| 8                                 | 7.439e-02         | 1.70  | 9.922e-02         | 1.27  | 7.982e+00       | 0.56  | 7.212e+00       | 0.78  |
| 16                                | 2.726e-02         | 1.45  | 3.650e-02         | 1.44  | 4.028e+00       | 0.99  | 3.757e+00       | 0.94  |
| 32                                | 7.790e-03         | 1.81  | 1.358e-02         | 1.43  | 1.834e+00       | 1.14  | 1.669e+00       | 1.17  |
| (Case 4: $x_E = V_E, x_e = b_e$ ) |                   |       |                   |       |                 |       |                 |       |
| $n$                               | PressErr( $C_E$ ) | Conv. | PressErr( $V_E$ ) | Conv. | VelErr( $c_e$ ) | Conv. | VelErr( $b_e$ ) | Conv. |
| 4                                 | 2.669e-01         | —     | 2.558e-01         | —     | 1.343e+01       | —     | 1.371e+01       | —     |
| 8                                 | 9.309e-02         | 1.52  | 8.639e-02         | 1.57  | 7.819e+00       | 0.78  | 7.085e+00       | 0.95  |
| 16                                | 3.738e-02         | 1.32  | 2.855e-02         | 1.60  | 3.827e+00       | 1.03  | 3.587e+00       | 0.98  |
| 32                                | 1.355e-02         | 1.46  | 7.926e-03         | 1.85  | 1.704e+00       | 1.17  | 1.572e+00       | 1.19  |

Table 4: Convergence rates for the four test cases with random Voronoi meshes in three-dimensions. We consistently find at least first-order convergence for both pressure and velocity. Also note that the pressure solution is most accurate when the error is calculated at the  $x_E$  point, and a slightly smaller error for the velocity when calculating the error at point  $x_e$ .

Special care must be taken when solving for porous media multi-phase flow equations, as they can be sensitive to the kind of discretization used. Methods such as the two-point flux approximation (TPFA) have been well established for such equations. While the original MFD method definition encompasses TPFA over rectangular grids, it does not include TPFA for cases such as Voronoi meshes. We have demonstrated the connection between the MFD method and TPFA over Voronoi grids by defining a generalization of the original MFD method. Establishing this connection results in a reduction of the saddle-point system associated with the MFD method to a symmetric-positive definite system through a Schur’s complement. In the case of rectangular grids, we have observed that the generalization connects the MFD method with point-centered schemes.

We have presented a proof of stability and convergence of the generalization using tools from the MFE method and the standard MFD method. The analysis suggests that the modification made to the MFD method maintains first-order convergence for both the pressure and velocity unknowns. We have also demonstrated second-order convergence (or superconvergence) of the pressure unknown by employing special lifting operators. In the case of two-dimensional rectangular grids, we explicitly constructed such operators, which can be considered as shape functions for a traditional MFE method. The method corresponds to different choices for the test and trial  $H(\Omega; \text{div})$  spaces, or a Petrov-Galerkin type method.

## 6.1 Future Work

Further connections with methods from the Finite Volume literature should be investigated. This includes various MPFA methods [1] and point-centered schemes [21]. The method defined here might open further possibilities for finding discretization schemes that satisfy the discrete maximum principle [35]. In addition, we plan to extend the theoretical results of this work to include velocity superconvergence and construction of lifting operators for more element types.

## Acknowledgments

The authors acknowledge the financial support from the King Abdullah University of Science and Technology Academic Excellence Alliance, from Saudi Aramco and from the Center for Subsurface Modeling Industrial Affiliates. The second author was partially supported by DOE grant DE-FG02-04ER25617; the third author was partially supported by DOE grant DE-FG02-04ER25618 and NSF grant DMS 1418947. We also acknowledge helpful feedback from Todd Arbogast, Mojdeh Delshad, Leszek Demkowicz, Xin Yang and Guangri (Gary) Xue.

The visualization in this work was generated using Paraview scientific visualizer [25]. The code used to produce results was written in Python and uses the NumPy [46], SciPy [29] and Cython [7] libraries. The code uses sparse direct solvers found in spsolve library of SciPy. For larger models, the code has been coupled with PETSc [6] using the petsc4py interface [17]. For cell volume and centroid computations, the code uses the algorithm defined in [38]. The Voronoi diagrams were generated using the Voro++ software package [43].

## References

- [1] I. Aavatsmark. An introduction to multipoint flux approximations for quadrilateral grids. *Computational Geosciences*, 6(3-4):405–432, 2002.
- [2] S. Agmon. *Lectures on elliptic boundary value problems*, volume 2. American Mathematical Soc., 2010.
- [3] T. Arbogast, M. Wheeler, and I. Yotov. Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences. *SIAM Journal on Numerical Analysis*, 34(2):pp. 828–852, 1997.
- [4] T. Arbogast, C. Dawson, P. Keenan, M. Wheeler, and I. Yotov. Enhanced cell-centered finite differences for elliptic equations on general geometry. *SIAM Journal on Scientific Computing*, 19(2):404–425, 1998.
- [5] K. Aziz. Reservoir simulation grids: Opportunities and problems. *Journal of Petroleum Technology*, pages 658–663, 1993.
- [6] S. Balay, M. Adams, J. Brown, P. Brune, K. Buschelman, V. Eijkhout, W.D. Gropp, D. Kaushik, M. Knepley, L. McInnes, K. Rupp, B. Smith, and H. Zhang. PETSc Web page. <http://www.mcs.anl.gov/petsc>, 2014.
- [7] S. Behnel, R. Bradshaw, C. Citro, L. Dalcin, D. Seljebotn, and K. Smith. Cython: The best of both worlds. *Computing in Science and Engineering*, 13(2):31–39, 2011.
- [8] S. Brenner and L. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2nd edition, April 2002.
- [9] F. Brezzi and A. Buffa. Innovative mimetic discretizations for electromagnetic problems. *Journal of Computational and Applied Mathematics*, 234(6):1980 – 1987, 2010.
- [10] F. Brezzi, A. Buffa, and K. Lipnikov. Mimetic finite differences for elliptic problems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 43(02):277–295, 2009.
- [11] F. Brezzi, J. Douglas Jr, and L. Marini. Two families of mixed finite elements for second order elliptic problems. *Numerische Mathematik*, 47(2):217–235, 1985.
- [12] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. New York: Springer-Verlag, 1991.
- [13] F. Brezzi, K. Lipnikov, and M. Shashkov. Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.*, 43(5):1872–1896 (electronic), 2005.
- [14] F. Brezzi, K. Lipnikov, and V. Simoncini. A family of mimetic finite difference methods on polygonal and polyhedral meshes. *Mathematical Models and Methods in Applied Sciences*, 15(10):1533–1551, 2005.

- [15] L. da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. Marini, and A. Russo. Basic principles of virtual element methods. *Mathematical Models and Methods in Applied Sciences*, 23(01):199–214, 2013.
- [16] L. da Veiga, V. Gyrya, K. Lipnikov, and G. Manzini. Mimetic finite difference method for the stokes problem on polygonal meshes. *Journal of Computational Physics*, 228(19):7215 – 7232, 2009.
- [17] L. Dalcin. petsc4py. <https://code.google.com/p/petsc4py/>, 2013.
- [18] J. Douglas Jr and J. Wang. Superconvergence of mixed finite element methods on rectangular domains. *Calcolo*, 26(2-4):121–133, 1989.
- [19] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Mathematical Models and Methods in Applied Sciences*, 20(02):265–295, 2010.
- [20] M. Edwards. Unstructured, control-volume distributed, full-tensor finite-volume schemes with flow based grids. *Computational Geosciences*, 6(3-4):433–452, 2002.
- [21] M. Edwards and C. Rogers. Finite volume discretization with imposed flux continuity for the general tensor pressure equation. *Computational Geosciences*, 2(4):259–290, 1998.
- [22] R. Ewing, R. Lazarov, and J. Wang. Superconvergence of the velocity along the gauss lines in mixed finite element methods. *SIAM journal on numerical analysis*, 28(4):1015–1029, 1991.
- [23] R. Eymard, T. Gallouët, and R. Herbin. *Finite volume methods*. Handbook of Numerical Analysis VII. North-Holland, Amsterdam, 2000.
- [24] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 69. SIAM, 2011.
- [25] A. Henderson. The paraview guide: A parallel visualization application. kitware inc. Technical report, 2007.
- [26] J. Hyman, M. Shashkov, and S. Steinberg. The numerical solution of diffusion problems in strongly heterogeneous non-isotropic materials. *Journal of Computational Physics*, 132(1):130–148, 1997.
- [27] M. Hyman and M. Shashkov. Mimetic discretizations for Maxwell’s equations. *Journal of Computational Physics*, 151(2):881 – 909, 1999.
- [28] R. Ingram, M. Wheeler, and I. Yotov. A multipoint flux mixed finite element method on hexahedra. *SIAM Journal on Numerical Analysis*, 48(4):1281–1312, 2010.
- [29] E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001.
- [30] R. Klausen and A. Stephansen. Convergence of multi-point flux approximation on general grids and media. *Inter. J. of Num. Analysis and Modeling*, 9(3):584–606, 2012.

- [31] R. Klausen and R. Winther. Convergence of multipoint flux approximations on quadrilateral grids. *Numerical methods for partial differential equations*, 22(6):1438–1454, 2006.
- [32] R. Klausen and R. Winther. Robust convergence of multi point flux approximation on rough grids. *Numerische Mathematik*, 104(3):317–337, 2006.
- [33] Y. Kuznetsov and S. Repin. New mixed finite element method on polygonal and polyhedral meshes. *Russian Journal of Numerical Analysis and Mathematical Modelling*, 18(3):261–278, 2003.
- [34] K. Lipnikov, G. Manzini, and M. Shashkov. Mimetic finite difference method. *Journal of Computational Physics*, 257, Part B(0):1163 – 1227, 2014. Physics-compatible numerical methods.
- [35] K. Lipnikov, G. Manzini, and D. Svyatskiy. Analysis of the monotonicity conditions in the mimetic finite difference method for elliptic problems. *Journal of Computational Physics*, 230(7):2620 – 2642, 2011.
- [36] K. Lipnikov, M. Shashkov, and D. Svyatskiy. The mimetic finite difference discretization of diffusion problem on unstructured polyhedral meshes. *Journal of Computational Physics*, 211:473–491, 2005.
- [37] K. Lipnikov, M. Shashkov, and I. Yotov. Local flux mimetic finite difference methods. *Numerische Mathematik*, 112:115–152, 2009.
- [38] B. Mirtich. Fast and accurate computation of polyhedral mass properties. *journal of graphics tools*, 1(2):31–50, 1996.
- [39] M. Nakata, A. Weiser, and M. Wheeler. Some superconvergence results for mixed finite element methods for elliptic problems on rectangular domains. *The Mathematics of Finite Elements and Applications V, JR Whiteman, ed., Academic Press, London*, pages 367–389, 1985.
- [40] C. Palagi and K. Aziz. Use of Voronoi grid in reservoir simulation. *SPE Advanced Technology Series*, 2:69–77, 1994.
- [41] P. Raviart and J. Thomas. A mixed finite element method for 2-nd order elliptic problems. In *Mathematical aspects of finite element methods*, pages 292–315. Springer, 1977.
- [42] T. Russell and M. Wheeler. Finite element and finite difference methods for continuous flows in porous media. In R.E. Ewing, editor, *Mathematics of Reservoir Simulation, Frontiers in Applied Mathematics Vol. 1*. SIAM, 1983.
- [43] C. Rycroft. Voro++: A three-dimensional voronoi cell library in c++. 2009.
- [44] A. Samarskii, V. Tishkin, A. Favorskii, and M. Shashkov. Use of the support operator method for constructing difference analogues of operations of tensor analysis. *Differentsial nye Uravneniya*, 18(7):1251–1256, 1287, 1982. English translation: *Differential Equations* 18 (1982), no. 7, 881–885.

- [45] M. Shashkov. *Conservative finite-difference methods on general grids*. CRC Press, Boca Raton, FL, 1996.
- [46] S. Van Der Walt, S. Colbert, and G. Varoquaux. The numpy array: a structure for efficient numerical computation. *Computing in Science and Engineering*, 13(2):22–30, 2011.
- [47] G. Vitaliy and K. Lipnikov. High-order mimetic finite difference method for diffusion problems on polygonal meshes. *Journal of Computational Physics*, 227:8841–8854, October 2008.
- [48] A. Weiser and M. Wheeler. On convergence of block-centered finite differences for elliptic problems. *SIAM Journal on Numerical Analysis*, 25(2):351–375, 1988.
- [49] M. Wheeler, G. Xue, and I. Yotov. A multipoint flux mixed finite element method on distorted quadrilaterals and hexahedra. *Numerische Mathematik*, 121(1):165–204, 2012.
- [50] M. Wheeler and I. Yotov. A multipoint flux mixed finite element method. *SIAM Journal on Numerical Analysis*, 44(5):2082–2106, 2006.