J. KSIAM **Vol.5**, No.1, 55-78, 2001

Comparing Two Approaches of Analyzing Mixed Finite Volume Methods

So-Hsiang Chou and Shengrong Tang

Abstract

Given the anisotropic Poisson equation $-\nabla \cdot \mathcal{K}\nabla p = f$, one can convert it into a system of two first order PDEs: the Darcy law for the flux $\mathbf{u} = -\mathcal{K}\nabla p$ and conservation of mass $\nabla \cdot \mathbf{u} = f$. A very natural mixed finite volume method for this system is to seek the pressure in the nonconforming P1 space and the Darcy velocity in the lowest order Raviart-Thomas space. The equations for these variables are obtained by integrating the two first order systems over the triangular volumes. In this paper we show that such a method is really a standard finite element method with local recovery of the flux in disguise. As a consequence, we compare two approaches in analyzing finite volume methods (FVM) and shed light on the proper way of analyzing non co–volume type of FVM. Numerical results for Dirichlet and Neumann problems are included.

1 Introduction

Consider the variable-coefficient Poisson equation in a polygonal domain $\Omega \subset \mathcal{R}^2$

$$\begin{cases} -\nabla \cdot \mathcal{K}\nabla p = f & \text{in } \Omega, \\ p = 0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $\mathcal{K} = \mathcal{K}(\mathbf{x})$ is a symmetric positive definite matrix function such that there exist two positive constants α_1 and α_2 with

$$\alpha_1 \xi^T \xi \le \xi^T \mathcal{K}(\mathbf{x}) \xi \le \alpha_2 \xi^T \xi \qquad \forall \xi \in \mathcal{R}^2, \mathbf{x} \in \bar{\Omega}. \tag{2}$$

Now let us introduce a flux variable $\mathbf{u} := -\mathcal{K}\nabla p$ and write the above equation as the system of first order partial differential equations

$$\begin{cases} \nabla \cdot \mathbf{u} - f = 0 & \text{in } \Omega \\ \mathbf{u} + \mathcal{K}\nabla p = 0 & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega \end{cases} \tag{3}$$

AMS subject classifications. 65F10, 65N20, 65N30
Key words and phrases: mixed finite elements, finite volume, box method, anisotropic porous media, $P1$ nonconforming, mixed hybrid method

Problems (1) and (3) are equivalent when the right hand side f, the diffusion tensor $\mathcal{K}$ and the domain Ω (e.g. convex) are such that the solution is smooth enough, e.g., $\mathbf{u} \in H^1(\Omega)^2, p \in H_0^1(\Omega) \cap H^2(\Omega)$.

This system can be interpreted as modeling an incompressible single phase flow in a reservoir; ignoring gravitational effects. The matrix $\mathcal{K}$ is the mobility κ/μ, the ratio of permeability tensor to viscosity of the fluid, $\mathbf{u}$ is the Darcy velocity and p the pressure. The second equation is the Darcy law and the first represents conservation of mass with f standing for a source or sink term. Since κ is in general discontinuous due to different rock formations, separating the Darcy law from the second order equation and discretizing it directly together with the mass conservation may lead to a better numerical treatment on the velocity than just computing it from the pressure via the Darcy law. This approach is well known in the finite element circle [20], but the same approach can be applied in conjunction with the finite volume method (termed mixed finite volume methods) as well [6, 9, 10, 11, 12, 20, 21]. For other similarly related issues, see also [17, 18, 22].

Let $\mathcal{T}_h = \{K_j\}_{j=1}^{NT}$ be the usual non–overlapping finite element triangulation of the domain $\Omega = \cup_{K\in\mathcal{T}_h} K$. Furthermore $\mathcal{T}_h$ is assumed to be regular, that is, $\min_{K\in\mathcal{T}_h} d(K)/\rho(K) \geq C$ for a constant C independent of h. Here $\rho(K)$ is the diameter of triangle K; $d(K)$ the diameter of the inscribed circle of K, and $h = \max_{K\in\mathcal{T}_h} \rho(K)$. We denote the area of K by $|K|$, by $\mathcal{A} = \mathcal{A}_i \cup \mathcal{A}_b$ the set of all edges of $\mathcal{T}_h$ consisting of the interior edge set $\mathcal{A}_i$ and boundary edge set $\mathcal{A}_b$. We use $N\mathcal{A}_i$ and $N\mathcal{A}_b$ to denote the number of interior edges and the number of boundary edges, respectively. The total number of edges is $N\mathcal{A} = N\mathcal{A}_i + N\mathcal{A}_b$.

Define the lowest order Raviart-Thomas space [19]

$$\mathbf{V}_h = \{\mathbf{u}_h \in H(\mathrm{div};\Omega) : \mathbf{u}_h|_K \in RT_0(K)\}$$

where $RT_0(K) = \{\mathbf{u} = (u^1, u^2) : u^1 = a + bx, u^2 = c + by \text{ in } K\}$ and the standard $P1$ nonconforming finite element space

$$Y_h = \{p_h|_K \in P_1(K) : p_h \text{ continuous at the middle point of each } e \in \partial K\}.$$

Consider approximating $\mathbf{u}$ by $\mathbf{u}_h \in \mathbf{V}_h$ and p by $p_h \in Y_h$ via a mixed finite volume approach on (3):

Find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Y_h$ such that

$$\begin{cases} (\nabla\cdot\mathbf{u}_h - f, \chi_K) &= 0 \quad \text{for all } K \in \mathcal{T}_h, \\ (\mathbf{u}_h + \mathcal{K}\nabla p_h, \underline{\chi}_K) &= 0 \quad \text{for all } K \in \mathcal{T}_h \end{cases} \tag{4}$$

where $p_h = 0$ at all midpoints of boundary edges. Here χ_K is the characteristic function of triangle K and $\underline{\chi}_K$ is any constant vector multiple of χ_K.

This method was introduced and analyzed for the case of $\mathcal{K} = \mathcal{I}$, the identity matrix, by Courbet & Croisille [15]. In this paper we study the general full tensor case. The full tensor case is important in porous media flow applications since it represents the more natural anisotropic case. Mathematically, the full tensor case is by no means a trivial extension of the simple Poisson equation, especially in the finite volume case: the presence of the full tensor often results in a nonsymmetric system and is harder to give physical interpretations [1]. Furthermore, unlike in the *mathematical analysis* of finite element methods, convergence of finite volume methods for the isotropic and anisotropic cases in general cannot be handled in a unified way. See [16] and the references therein for the difficulties involved. Also see [10, 11] where a uniform treatment is possible, but in these co-volume schemes Chou *et. al.* used two grid systems in discretization. It is therefore surprising that the extension of the Courbet & Croisille method has none of these drawbacks, having only one grid system for both variables.

This paper is the first in a series of two papers on comparing the analysis of the finite volume method from a finite element person's or a "pure" finite volume person's viewpoint. It grows out of the informal report [13]. We will present things in an elementary and non–terse way and the point will be how different viewpoints may lead to the same conclusions but in a roundabout way. The reader is referred to follow-up [14] for a more extensive and in-depth mathematical presentation.

First we adopt a pure finite volume person's point of view. In Sec. 2, we derive the discrete system, and in Sec. 3 we compute explicitly the element and global stiffness matrices whose proper interpretations lead to Thm 4.1, which says that the pressure approximation can be re–interpreted as the solution to the system of the standard $P1$ nonconforming finite elements applied to the second order elliptic problem (1). However, the process leading to that conclusion is long and technical. The advantage in this approach is that conservation laws are obvious from the start and the drawback is that the mathematical analysis is long and very uninspiring.

On the other hand, as shown in [8, 12, 11] it is fruitful to try to relate the analysis of a mixed finite volume method to a close finite element method. Adopting this viewpoint, we can prove Thm. 4.1 in a few short lines. The advantage of this approach is that mathematical analysis is neat, but conservation law is somewhat hidden. To keep the paper short, how to get the best of these two approaches is elucidated in the follow-up paper [14].

The remaining of this paper is organized as follows. In Sec. 5, we show the error estimates in p and $\mathbf{u}$. Finally, in the last section we provide numerical results for both Dirichlet and Neumann problems.

2 Problem formulation

Let us now represent the system (4) using proper basis functions. Notice that by divergence theorem, Eq. $(4)_1$ can be written as

$$\int_{\partial K} \mathbf{u}_h \cdot \mathbf{n} d\mathbf{x} - |K| f_K = 0 \tag{5}$$

where $f_K = \frac{1}{|K|}\int_K f d\mathbf{x}$ is the average of $f(\mathbf{x})$ over triangle K. Since ∇p_h is constant over K, Eq. $(4)_2$ can be written as

$$\int_K \mathbf{u}_h d\mathbf{x} + |K|\mathcal{A}_K \nabla p_h = 0 \tag{6}$$

where the matrix $\mathcal{A}_K = \frac{1}{|K|}\int_K \mathcal{K} d\mathbf{x}$ is the average of matrix $\mathcal{K}(\mathbf{x})$ over triangle K.

Figure 1: Local elements based on an interior edge $\mathbf{e}_1$.

With reference to Fig. 1, let λ_S be the usual nodal linear basis function associated with the vertex S of $K = K_L$. Recall that λ_S is one at S and zero at other two vertices and is also called barycentric or area coordinate function. For any $p_h(\mathbf{x}) \in Y_h$, we have the local representation on K

$$p_h(\mathbf{x})|_K = \sum_{e \in \partial K} p_e \varphi_e(\mathbf{x}) \tag{7}$$

where $\varphi_e(\mathbf{x}) = 1 - 2\lambda_S(\mathbf{x})$ is the local basis function of space Y_h on edge e with $\lambda_S(\mathbf{x})$ being barycentric coordinate of $\mathbf{x}$ with respect to vertex S opposite to e in triangle K. (Note that $e = e_1$ in Fig. 1.) It is easy to see $\nabla\varphi_e(\mathbf{x}) = \frac{|e|}{|K|}\mathbf{n}_e = const$, where $|e|$ is the length of edge e.

Given any triangular element $K \in \mathcal{T}_h$, we always orient K counterclockwise as shown in Fig. 1 (e.g. $K = K_L$ there). Then the three local basis functions associated with the three edges are as follows. For example, for the edge $e = e_1$ of $K = K_L$ in Fig. 1, we define $\mathbf{P}_{K,e}$

$$\mathbf{P}_{K,e}(\mathbf{x}) = \frac{1}{2|K|}\left[\begin{array}{c} x - x_S \\ y - y_S \end{array}\right] \qquad \forall \quad (x, y) \in K. \tag{8}$$

Note that

$$\mathbf{P}_{K,e}(\mathbf{x}) \cdot \mathbf{n} = \begin{cases} 1/|e| & \forall \quad \mathbf{x} \in e = S'S'' \\ 0 & \forall \quad \mathbf{x} \in SS' \\ 0 & \forall \quad \mathbf{x} \in SS''. \end{cases} \tag{9}$$

The other two basis functions $P_{K,e_i}, i = 2, 3$ are defined similarly.

For any $\mathbf{u}_h(\mathbf{x}) \in \mathbf{V}_h$, we have the local representation on K

$$\mathbf{u}_h(\mathbf{x})|_K = \sum_{e \in \partial K} u_e \mathbf{P}_{K,e}(\mathbf{x}) \tag{10}$$

where $u_e = \int_e \mathbf{u} \cdot \mathbf{n} ds$ is the flux across edge e. For any edge $a \in \mathcal{A}_i$ we define the global canonical basis of $\mathbf{V}_h$ associated with edge a as follows. If the edge a corresponds to edge $S'S''$ in the local ordering (cf. Fig. 1), then

$$\mathbf{P}_a(\mathbf{x}) = \mathbf{P}_{K_L, e}(\mathbf{x}) \chi_{K_L}(\mathbf{x}) - \mathbf{P}_{K_R, e}(\mathbf{x}) \chi_{K_R}(\mathbf{x}) \tag{11}$$

where a is oriented from K_L towards K_R. The global basis functions based on boundary edges are defined similarly.

Finally, by Taylor's expansion at the barycenter B of K we have on K

$$\mathbf{u}_h(\mathbf{x}) = \mathbf{u}_K + (\nabla \cdot \mathbf{u}_h)_K \mathbf{P}_K(\mathbf{x}) \tag{12}$$

where $\mathbf{u}_K = \frac{1}{|K|}\int_K \mathbf{u}_h d\mathbf{x} = -\mathcal{A}_K \nabla p_h$, the average of $\mathbf{u}_h(x)$ on K, and

$$\mathbf{P}_K(\mathbf{x}) = \frac{1}{2}\left[\begin{array}{c} x - x_B \\ y - y_B \end{array}\right] = \frac{|K|}{3} \sum_{e \in \partial K} \mathbf{P}_{K,e}(\mathbf{x}) \qquad \forall \quad (x, y) \in K \tag{13}$$

with (x_B, y_B) being the coordinates of B. Alternatively, one can also write

$$\mathbf{u}_h(\mathbf{x}) = -\mathcal{A}_K \nabla p_h + |K| f_K \mathbf{P}_K(\mathbf{x}). \tag{14}$$

Note that $|K| f_K \mathbf{P}_K(\mathbf{x})$ has zero mean on triangle K.

Let $\mathbf{u}_h \in \mathbf{V}_h$ and $p_h \in Y_h$ have the local representations (10) and (7), respectively. Then for each $K \in \mathcal{T}_h$, Eq. (5) implies

$$\sum_{e \in \partial K} u_e - |K| f_K = 0 \qquad \text{(NT equations)} \tag{15}$$

and Eq. (6) becomes

$$\sum_{e\in\partial K} u_e \int_K \mathbf{P}_{K,e}(\mathbf{x})d\mathbf{x} + p_e|K|\mathcal{A}_K\nabla\varphi_e(\mathbf{x}) = 0.$$

Define $\mathbf{Q}_e = \int_K \mathbf{P}_{K,e}(x)dx$, $\mathbf{N}_e = |e|\mathcal{A}_K\mathbf{n}_e$, and recall that $\nabla\varphi_e(\mathbf{x}) = \frac{|e|}{|K|}\mathbf{n}_e = const$ to get

$$\sum_{e\in\partial K} u_e\mathbf{Q}_e + p_e\mathbf{N}_e = 0 \qquad \text{(2NT equations)}. \tag{16}$$

Referring to Fig. 1, we see that by the one-point quadrature using the barycenter B

$$\mathbf{Q}_{e_1} = \mathbf{P}_{K,e_1}(B)|K| = 1/6(\overrightarrow{SS'} + \overrightarrow{SS''}) := 1/6(\mathbf{e}_3 - \mathbf{e}_2)$$

where $\overrightarrow{SS'} = \mathbf{e}_3$ and $\overrightarrow{S''S} = \mathbf{e}_2$. Hence

$$\sum_{e\in\partial K} \mathbf{Q}_e = 0. \tag{17}$$

Also note

$$\sum_{e\in\partial K} \mathbf{N}_e = \mathcal{A}_K(\sum_i |\mathbf{e}_i|\mathbf{n}_{e_i}) = 0. \tag{18}$$

On the boundary

$$p_a = 0 \qquad (N\mathcal{A}_b \text{ equations}). \tag{19}$$

Clearly

$$3NT + N\mathcal{A}_b = 2N\mathcal{A}. \tag{20}$$

and we have as many equations as unknowns: the number of unknowns $(u_a, p_a)_{a\in\mathcal{A}}$ being $2N\mathcal{A}$ and the total number of equations in (15), (16) and (19) being $3NT+N\mathcal{A}_b = 2N\mathcal{A}$.

Combining (15), (16) and (19), we see that system (4) becomes: Find $\mathbf{u}_h = \sum_{a\in\mathcal{A}} u_a\mathbf{P}_a(\mathbf{x})$, $p_h(\mathbf{x}) = \sum_{a\in\mathcal{A}} p_a\varphi_a(\mathbf{x})$ such that

$$\left\{ \begin{array}{ccll} \sum_{e\in\partial K} u_e & = & |K|f_K & \forall K \in \mathcal{T}_h, \\ \sum_{e\in\partial K}(u_e\mathbf{Q}_e + p_e\mathbf{N}_e) & = & 0 & \forall K \in \mathcal{T}_h, \\ p_a & = & 0 & \forall a \in \mathcal{A}_b. \end{array} \right. \tag{21}$$

A remark about notation is in order here. We emphasize that the notation $u_a, a \in \mathcal{A}$ is reserved for the component with respect to the global basis whereas the notation $u_{K,e}, e \in \partial K$ is the component with respect to the local basis. When there is no danger of confusion we simply use u_e instead of $u_{K,e}$. Later in sec. 5, we shall show the existence and uniqueness of the above system.

3 Element and global stiffness matrices

In this section we shall work out the details of implementation of the resulting discrete systems.

Denote by $U = (u_a)_{a\in\mathcal{A}}$ the global vector of $\mathbf{u}_h(\mathbf{x})$ onto basis $\{\mathbf{P}_a(\mathbf{x})\}$ and $P = (p_a)_{a\in\mathcal{A}}$ the global vector of $p_h(\mathbf{x})$ onto basis $\{\varphi_a(\mathbf{x})\}$. Also define U_K and P_K to be the local vectors of $\mathbf{u}_h(\mathbf{x})$ and $p_h(\mathbf{x})$ onto local basis functions in the triangle K

$$\begin{aligned} U_K &= [u_{e_1}, u_{e_2}, u_{e_3}]^T, && (22) \\ P_K &= [p_{e_1}, p_{e_2}, p_{e_3}]^T && (23) \end{aligned}$$

where e_1, e_2 and e_3 are three edges of K.

Then system (21) can be written as:

$$\bar{L}_K U_K + \bar{M}_K P_K = \bar{F}_K \qquad \forall K \in \mathcal{T}_h \tag{24}$$

where $\bar{L}_K, \bar{M}_K \in \mathcal{R}^{3\times 3}, \bar{F}_K \in \mathcal{R}^3$ with

$$\bar{L}_K = \begin{bmatrix} 1 & 1 & 1 \\ \mathbf{Q}_{e_1} & \mathbf{Q}_{e_2} & \mathbf{Q}_{e_3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ \frac{1}{6}(\mathbf{e}_3 - \mathbf{e}_2) & \frac{1}{6}(\mathbf{e}_1 - \mathbf{e}_3) & \frac{1}{6}(\mathbf{e}_2 - \mathbf{e}_1) \end{bmatrix},$$

$$\bar{M}_K = \begin{bmatrix} 0 & 0 & 0 \\ \mathbf{N}_{e_1} & \mathbf{N}_{e_2} & \mathbf{N}_{e_3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ |e_1|\mathcal{A}_K\mathbf{n}_{e_1} & |e_2|\mathcal{A}_K\mathbf{n}_{e_2} & |e_3|\mathcal{A}_K\mathbf{n}_{e_3} \end{bmatrix},$$

$$\bar{F}_K = \begin{bmatrix} |K| f_K \\ 0 \\ 0 \end{bmatrix}.$$

We know that the matrix $\bar{L}_K$ is nonsingular by (17) and hence (24) can be rewritten as

$$U_K = F_K - M_K P_K \qquad \forall K \in \mathcal{T}_h \tag{25}$$

where $M_K = \bar{L}_K^{-1}\bar{M}_K$ and $F_K = \bar{L}_K^{-1}\bar{F}_K$. It is easy to check

$$F_K = \frac{|K|}{3} f_K \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{26}$$

We can eliminate the unknowns $(u_a)_{a\in\mathcal{A}}$ to obtain a system in the pressures alone. If a is an interior edge with orientation from $K_L(a)$ towards $K_R(a)$, $a = e_L$ in $K_L(a)$, $a = e_R$ in $K_R(a)$, then the continuity of the flux across a gives the identity $U_{K_L,e_L} = -U_{K_R,e_R}$ holds. (Here U_{K_L,e_L} means the component of U_{K_L} in (22) corresponding to the edge e_L.) Thus we have the (scalar) identity

$$[M_{K_L} P_{K_L}]_{e_L} + [M_{K_R} P_{K_R}]_{e_R} = F_{K_L,e_L} + F_{K_R,e_R} \qquad \forall a \in \mathcal{A}_i. \tag{27}$$

This assembly along with $p_a = 0$ for all $a \in \mathcal{A}_b$ then gives rise to an $N\mathcal{A} \times N\mathcal{A}$ linear system in the unknown $P = (p_a)_{a\in\mathcal{A}}$

$$\mathcal{M}P = \mathcal{F}. \tag{28}$$

We denote $P_i \in \mathcal{R}^{N\mathcal{A}_i}$, $\mathcal{F}_i \in \mathcal{R}^{NA_i}$, and $\mathcal{M}_i \in \mathcal{R}^{NA_i \times NA_i}$ the sub-vectors or submatrix of $P, \mathcal{F}$ and $\mathcal{M}$ corresponding to interior edge set $\mathcal{A}_i$. The resulting system is $\mathcal{M}_i P_i = \mathcal{F}_i$. Next, we will give a simple form of local stencil M_K which is a symmetric positive matrix. Then we will introduce a way to assemble global stiffness matrix $\mathcal{M}_i$ from local stencil M_K. Finally we show that the matrix $\mathcal{M}_i$ is symmetric positive definite.

Lemma 3.1 The local stencil M_K is symmetric positive semi-definite.

Proof 1 Let us compute $\bar{L}_K^{-T}$ first. Referring to the notation in Fig. 1, we have

$$\bar{L}_K^T = \begin{bmatrix} 1 & 1/6(\mathbf{e}_3 - \mathbf{e}_2)^T \\ 1 & 1/6(\mathbf{e}_1 - \mathbf{e}_3)^T \\ 1 & 1/6(\mathbf{e}_2 - \mathbf{e}_1)^T \end{bmatrix}.$$

Using the fact $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 = 0$ and elementary geometry, we can easily see that

$$\bar{L}_K^{-T} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{|\mathbf{e}_1|}{|K|}\mathbf{n}_{\mathbf{e}_1} & \frac{|\mathbf{e}_2|}{|K|}\mathbf{n}_{\mathbf{e}_2} & \frac{|\mathbf{e}_3|}{|K|}\mathbf{n}_{e_3} \end{bmatrix}$$

where $\mathbf{n}_{e_i}$ are the unit outward normal to $\mathbf{e}_i$. Thus $M_K = \bar{L}_K^{-1}\bar{M}_K$ has entries

$$(M_K)_{ij} = \frac{|\mathbf{e}_i||\mathbf{e}_j|}{|K|}\mathbf{n}_{\mathbf{e}_i}^T \mathcal{A}_K \mathbf{n}_{\mathbf{e}_j} = \frac{|\mathbf{e}_i||\mathbf{e}_j|}{|K|}\mathbf{n}_{\mathbf{e}_j}^T \mathcal{A}_K \mathbf{n}_{\mathbf{e}_i} = (M_K)_{ji}$$

and hence M_K is symmetric.

Now let $R(\theta)$ be the rotation matrix through an angle of θ, then $\mathbf{n}_{e_i} = R(-\frac{\pi}{2})\mathbf{e}_i/|\mathbf{e}_i|$ and so

$$(M_K)_{ij} = \frac{1}{|K|}{\mathbf{e}_i}^T \tilde{\mathcal{A}}_K \mathbf{e}_j \tag{29}$$

where

$$\tilde{\mathcal{A}}_K = \begin{bmatrix} c & -b \\ -b & a \end{bmatrix} \qquad if \quad \mathcal{A}_K = \begin{bmatrix} a & b \\ b & c \end{bmatrix}. \tag{30}$$

Next we show M_K is positive semi-definite. Denote $E = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$, then

$$M_K = \frac{1}{|K|}E^T \tilde{\mathcal{A}}_K E. \tag{31}$$

and

$$\begin{aligned}\xi^T M_K \xi &= \frac{1}{|K|}\xi^T E^T \tilde{\mathcal{A}}_K E\xi = \frac{1}{|K|}\xi^T E^T R(-\frac{\pi}{2})^T \mathcal{A}_K R(-\frac{\pi}{2})E\xi \\ &\geq \alpha_1 \frac{1}{|K|}|R(-\frac{\pi}{2})E\xi|^2 = \alpha_1 \frac{1}{|K|}|E\xi|^2.\end{aligned}$$

Therefore M_K is positive semi-definite.

Remark: Note if we take $\mathcal{A}_K = I$, from (29) a simple computation shows that

$$M_K = 2\begin{bmatrix} d_2+d_3 & -d_3 & -d_2 \\ -d_3 & d_1+d_3 & -d_1 \\ -d_2 & -d_1 & d_1+d_2 \end{bmatrix}$$

where $d_i = \cot\theta_i$, where θ_i is the angle opposite to the edge e_i. It is 4 times the local stencil in the standard FVM for Poisson problem. (see [3]) and same as that of the mixed covolume scheme. (see [14])

Now let $\mathcal{I} := \{1, 2, \ldots, N\mathcal{A}_i\}$ be a global ordering of the interior edges in $\mathcal{A}_i$, and for $l = 1, \ldots, NT$, let $\{1, 2, 3\}$ be a local ordering of the edges of triangular element $K^{(l)}$. We use $g_j^{(l)}$ to denote the global edge number of the edge in element $K^{(l)}$ that has local edge number j. Also we use the notation $M^{(l)}$ for the local stencil M_K when $K = K^{(l)}$. Suppose that $a \in \mathcal{A}_i$ is the intersection of two elements $K^{(l)}$ and $K^{(m)}$ with orientation from $K^{(l)}$ to $K^{(m)}$. Let $a = e_s^{(m)}$ on $K^{(m)}$ and $a = e_t^{(l)}$ where s and t are the local edge numbers. Now let $K_L = K^{(l)}$ and $K_R = K^{(m)}$ in (27), then its left hand side can be expressed as

$$LHS = \sum_{j=1}^{3} m_{tj}^{(l)} p_{g_j^{(l)}} + \sum_{j=1}^{3} m_{sj}^{(m)} p_{g_j^{(m)}} \tag{32}$$

in terms of pressures $p_r, r \in \mathcal{I}$ (globally indexed) and the entries $m_{tj}^{(l)}$ of the three by three matrix $M^{(l)}$ and so on. The above suggests we define a global matrix for each element as follows. For each l, $1 \leq l \leq NT$ define the matrix $\hat{M}^{(l)} = \hat{M}_K \in \mathcal{R}^{N\mathcal{A}_i \times N\mathcal{A}_i}$ associated with $K = K^{(l)}$ so that $\hat{m}_{ij}^{(l)} = 0$ if the (global) edges i and j are not in $K^{(l)}$. Otherwise we set $\hat{m}_{ij}^{(l)} = m_{st}^{(l)}$ where $i = g_s^{(l)}$ and $j = g_t^{(l)}$. Obviously, $\hat{M}_K$ is a symmetric matrix since M_K is.

We can now write (32) in terms of this new matrix and global indices. It is then not hard to conclude that

$$\mathcal{M}_i = \textstyle\sum_{K\in\mathcal{T}_h} \hat{M}_K, \tag{33}$$

$$\mathcal{F}_i = \textstyle\sum_{K\in\mathcal{T}_h} \hat{F}_K. \tag{34}$$

In fact, let $\alpha \in \mathcal{I}$ and the edge $e_\alpha = a = K^{(l)} \cap K^{(m)}$. then

$$\begin{aligned}
\left(\sum_{K\in\mathcal{T}_h} \hat{M}_K P_i\right)_\alpha &= \sum_{\beta=1}^{N\mathcal{A}_i}\sum_{k=1}^{NT} \hat{m}^{(k)}_{\alpha\beta}(p_i)_\beta \\
&= \sum_{\beta=1}^{N\mathcal{A}_i} \left\{ \hat{m}^{(l)}_{\alpha\beta}(p_i)_\beta + \hat{m}^{(m)}_{\alpha\beta}(p_i)_\beta \right\} \\
&= \sum_{j=1}^{3} m^{(l)}_{tj}(p_i)_{g_j^{(l)}} + m^{(m)}_{sj}(p_i)_{g_j^{(m)}} \\
&= (\mathcal{M}_i P_i)_\alpha. \qquad (35)
\end{aligned}$$

Hence $\mathcal{M}_i$ is symmetric. Moreover

$$P_i^T \mathcal{M}_i P_i = \sum_{K\in\mathcal{T}_h} P_i^T \hat{M}_K P_i = \sum_{K\in\mathcal{T}_h} P_K^T M_K P_K. \qquad (36)$$

So from Lemma (3.1), we have $P_i^T \mathcal{M}_i P_i \geq 0$. Later in Theorem 5.2 we will show the uniqueness of solution for the system (21). Therefore we obtain the following theorem:

Theorem 3.1 *The global stiffness matrix $\mathcal{M}_i$ corresponding to the interior edge set $\mathcal{A}_i$ is symmetric positive definite.*

Algorithm:
do for each $K \in \mathcal{T}_h$
    evaluate $\tilde{\mathcal{A}}_K$, f_K
    evaluate M_K by (29), F_K by (26)
    assemble M_K to $\mathcal{M}_i, F_K$ to $\mathcal{F}_i$
enddo
solve $\mathcal{M}_i P_i = \mathcal{F}_i$
evaluate U from (25), with boundary condition.

4 a Mixed box method is a FEM plus local flux recovery

In this section we first prove in Thm. 4.1 equivalence between the pressure approximation in our box method and a $P1$ nonconforming finite element method applied to the elliptic problem (1) with a *modified* right hand side. Then we show this leads to a better understanding of the box method. It turns out that the mixed box method is nothing but a standard FEM with accurate local recovery of the flux. This is the content of Thm. 4.2 with zero absorption.

Throughout the rest of this paper, we use the standard notation $W^{1,p}$ for the usual Sobolev spaces and $|\cdot|_{m,K}$, $\|\cdot\|_{m,K}$ for the semi and full H^m–norm, $m = 0, 1, 2$. We omit the subscript K when $K = \Omega$ and sometimes use $\|\cdot\|$ when writing an L^2 norm. Also $|\cdot|_{H(\text{div};\Omega)}$ is the $H(\text{div};\Omega)$ semi-norm.

We now show the equivalence theorem.

Theorem 4.1 The linear system (28) is the same as the discrete system resulting from the standard $P1$ nonconforming FEM : Find $p_h \in Y_{h,0}$ such that

$$a_h(p_h, q_h) = (f_h, q_h) \qquad \forall q_h \in Y_{h,0} \tag{37}$$

where $f_h = P_h f$ is the L_2 projection to the piecewise constant space L_h.

Proof 2 It suffices to show the two methods have the same element stiffness matrix and the same right hand side. The element stiffness matrix associated with element K from (41) can be obtained as follows. Let $\varphi_i = 1 - 2\lambda_i$, $i = 1, 2, 3$ with λ_i being the barycentric coordinates. Noting that $\nabla\varphi_i = \frac{|{}^{\mathbf{M}}\mathbf{e}_i|}{|K|}\mathbf{n}_i$, we have

$$\begin{aligned}
\int_K (\mathcal{K}\nabla\varphi_i)\cdot\nabla\varphi_j d\mathbf{x} &= \int_K (\mathcal{K}\frac{|\mathbf{e}_i|}{|K|}\mathbf{n}_i)\cdot\frac{|\mathbf{e}_j|}{|K|}\mathbf{n}_j d\mathbf{x} \\
&= \frac{|\mathbf{e}_i|}{|K|}\mathbf{n}_i^T(\int_K \mathcal{K}d\mathbf{x})\frac{|\mathbf{e}_j|}{|K|}\mathbf{n}_j \\
&= \frac{1}{|K|}|\mathbf{e}_i|\mathbf{n}_i^T\mathcal{A}_K|\mathbf{e}_j|\mathbf{n}_j \\
&= \frac{1}{|K|}\mathbf{e}_i^T\tilde{\mathcal{A}}_K\mathbf{e}_j
\end{aligned}$$

which is exactly (29).

Since $f_h = P_h f$, the L_2 projection, $f_h|_K = \frac{1}{|K|}\int_K f$. Thus

$$\begin{aligned}
(f_h, \varphi_j) &= \int_K f_h\varphi_j dx \\
&= \frac{f_h|_K\,|K|}{3} \\
&= \frac{|K|f_K}{3}
\end{aligned}$$

which is exactly (26). This completes the proof.

Remark. We must emphasize this theorem is proved because of the *hard* work in the previous sections. Our line of thinking follows that of [15]. In other words, we try to generalize their method to the anisotropic case, but our crucial new observation is the above theorem. It says the mixed *finite volume* method is related to a *finite element* method. The next theorem says the hard work in the previous sections can be avoided altogether if we have a change of viewpoint.

To bring out an important difference between the mixed box method(finite volume) method and the finite element method, we will do more and actually show the equivalence theorem for the original problem (1) with an added absorption term i.e.,

$$\begin{cases} -\nabla\cdot\mathcal{K}\nabla p+\alpha_0 p &= f \qquad \text{in } \Omega, \\ p &= 0 \qquad \text{on } \partial\Omega, \end{cases} , \tag{38}$$

where α_0 is a nonnegative piecewise constant function with respect to $\mathcal{T}_h$. The associated weak formulation is to find $p \in H_0^1$ such that

$$a(p,q)=(f,q) \qquad \forall q\in H_0^1 \tag{39}$$

where

$$a(p,q):=\int_\Omega (\mathcal{K}\nabla p)\cdot\nabla q+\alpha_0 pq d\mathbf{x}. \tag{40}$$

Let

$$Y_{h,0} \;:=\; \{q\in L^2(\Omega): q|_K\in P_1(K), \forall K\in\mathcal{T}_h; q \text{ is continuous at the midpoints of interior edges and vanishes at the midpoints of boundary edges}\}.$$

The standard $P1$ nonconforming FEM discretization is : Find $\tilde{p}_h \in Y_{h,0}$ such that

$$a_h(\tilde{p}_h,q_h)=(f,q_h) \qquad \forall q_h\in Y_{h,0} \tag{41}$$

where

$$a_h(\tilde{p}_h,q_h):=\sum_{j=1}^{NT}(\mathcal{K}\nabla\tilde{p}_h,\nabla q_h)_{K_j}+(\alpha_0 p_h,q_h)_{K_j}, \tag{42}$$

where $(\cdot,\cdot)_K$ is the L^2 inner product on K. Define the semi-norm

$$|q|_h:=(\sum_{K\in\mathcal{T}_h}|q|_{1,K}^2)^{1/2} \qquad \forall q\in H_0^1\oplus Y_{h,0}. \tag{43}$$

It is clear $|\cdot|_h$ is a full norm on space $Y_{h,0}$ in Dirichlet Case. It is well known [7, 4, 5] that the solution p_h of system (41) converges to solution p of system (39) : there exists a constant C independent of h such that

$$\|p-\tilde{p}_h\|_0+h|p-\tilde{p}_h|_h\le Ch^2\|p\|_2 \tag{44}$$

provided that the problem data is smooth enough so that the elliptic regularity estimate $\|p\|_2 \le C\|f\|_0$ holds. For example, if $f \in L^2$ and $\mathcal{K} \in C^1(\bar{\Omega})$ on a convex domain Ω, then $p \in H^2$ is guaranteed. (See p. 4 of [4] and the references therein.) We now show the equivalence theorem.

Theorem 4.2 Consider the problem of finding $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Y_{h,0}$ such that

$$(\mathcal{K}\nabla p_h, \nabla q_h) + \sum_K \alpha_K (p_K, q_h)_K = (f_h, q_h) \qquad \forall q_h \in Y_{h,0} \tag{45}$$

and over each K in $\mathcal{T}_h$

$$\mathbf{u}_h = -\mathcal{A}_K \nabla p_h + (f_K - \alpha_K p_K)\mathbf{P}_K \tag{46}$$

where $f_h = P_h f$ is the L_2 projection to the piecewise constant space L_h and $p_K = \frac{1}{|K|}\int_K p_h dx, f_K = \frac{1}{|K|}\int_K f dx = f_h|_K$ along with the mixed box method of finding $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Y_{h,0}$ such that

$$(\nabla \cdot \mathbf{u}_h + \alpha_0 p_h - f, \chi_K) = 0, \tag{47}$$

$$(\mathbf{u}_h + \mathcal{K}\nabla p_h, \underline{\chi}_K) = 0. \tag{48}$$

Then the two above problems are equivalent.

Proof 3 We first show that (45)–(46) implies (47)–(48).

Take divergence on (46), recall (13), and integrate against the characteristic function χ_K to see (46) implies (47). Now integrate (46) against χ_K and use the fact $(\mathbf{P}_K, \underline{\chi}_K) = 0$(one point quadrature rule using the barycenter B) to get (48).

Secondly, we prove that (47)–(48) implies (45)–(46).

¿From (47) we see that on K

$$\nabla \cdot \mathbf{u}_h = f_K - \alpha_K p_K.$$

Since $\mathbf{u}_h \in \mathbf{V}_h$, by Taylor's expansion

$$\mathbf{u}_h = \mathbf{u}_K + \frac{1}{2}\nabla \cdot \mathbf{u}_h(\mathbf{x} - \mathbf{x}_B),$$

where $\mathbf{u}_K = \frac{1}{|K|}\int_K \mathbf{u}_h d\mathbf{x} = -\mathcal{A}_K \nabla p_h$ by (48). So

$$\mathbf{u}_h = -\mathcal{A}_K \nabla p_h + (f_K - \alpha_K p_K)\mathbf{P}_K$$

which is (46). On the other hand by (48), integration by parts and (47)

$$\begin{aligned}\sum_K (\mathcal{K}\nabla p_h, \nabla q_h)_K &= \sum_K (-\mathbf{u}_h, \nabla q_h)_K \\ &= \sum_K (\nabla \cdot \mathbf{u}_h, q_h)_K - (\mathbf{u}_h \cdot \mathbf{n}, q_h)_{\partial K} \\ &= \sum_K (f_K - \alpha_K p_K, q_h)\end{aligned}$$

which is (45). Notice that the $(\cdot,\cdot)_{\partial K}$ terms cancelled upon summation since $\mathbf{u}_h \cdot \mathbf{n}$ and q_h are continuous at midpoints of edges.

Remark. One notes that the mixed finite volume method is equivalent to the standard nonconforming method (with a modified right hand side) when there is no absorption term ($\alpha_0 = 0$). In the presence of the absorption term, the mixed method is equivalent to a nonconforming method with the absorption term evaluated by a one-point quadrature at the element barycenter. The approximate p_h so produced by (45) generates automatically a flux with continuous normal components across edges via (46). In other words, *the usual nonconforming method does not produce continuous flux unless* $\alpha_0 = 0$.

5 Error estimates

In this section we will prove the existence and uniqueness of solution for the system (21) and some error estimates in an energy norm. Through this section the letter C denotes a generic positive constant, independent of h and not necessarily the same in each occurrence. Let us define two energy norms which by (2) are equivalent:

$$|q|_h^2 = \sum_K |q|_{h,K}^2 := \sum_K \int_K |\nabla q|^2 d\mathbf{x} \quad \forall q \in H^1(\Omega) \oplus Y_h.$$

$$|q|_E^2 = \sum_K |q|_{E,K}^2 := \sum_K \int_K \nabla q^T \mathcal{K}(\mathbf{x}) \nabla q d\mathbf{x} \quad \forall q \in H^1(\Omega) \oplus Y_h.$$

The next lemma shows they are actually full norms on the space $Y_{h,0} = \{q_h \in Y_h, q_h = 0$ at all midpoints of boundary edges$\}$.

Lemma 5.1 The discrete energy semi-norm $|q_h|_h$ is a norm on the space $Y_{h,0}$.

Proof 4 Let $q_h \in Y_{h,0}$ such that $|q_h|_h = 0$. The gradient of q_h is zero in each cell $K \in \mathcal{T}_h$. Hence q_h is constant in each cell K. Since q_h is continuous at the middle point of each edge e of $\mathcal{T}_h$ and $q_h = 0$ on $\partial\Omega$, we have $q_h = 0$ in Ω.

The next theorem shows the existence and uniqueness of solution for system (21).

Theorem 5.1 The discrete system (21) has a unique solution $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Y_{h,0}$

Proof 5 *Since the number of unknowns equals the number of equations in (21) we only need to show that* $\mathbf{u}_h = \mathbf{0}$ *and* $p_h = 0$ *when* $f = 0$.

By Thm. 4.1 or 4.2 and the preceding lemma $p_h = 0$ *when* $f = 0$. *On the other hand,* $\mathbf{u}_h = \mathbf{u}_K = -\mathcal{A}_K \nabla p_h = 0$.

Also we have the following stability condition:

Lemma 5.2 If $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Y_{h,0}$ is the solution of system (21), then there exist positive constants C_1 and C_2, independent of h, such that

$$C_1|p_h|_h \leq \|\mathbf{u}_h\|_{0,\Omega} \leq C_2(|p_h|_h + h\|f\|_{0,\Omega}) \tag{49}$$

$$|\mathbf{u}_h|_h := \left(\sum_K \|\nabla \mathbf{u}_h\|_{0,K}^2\right)^{1/2} \leq \frac{1}{\sqrt{2}}\|f\|_{0,\Omega} \tag{50}$$

Proof 6 We first show (50). From (14) $\nabla \mathbf{u}_h = |K| f_K \nabla \mathbf{P}_K(\mathbf{x})$ and from (13) $\|\nabla \mathbf{P}_K\|_{0,K}^2 = \frac{1}{2|K|}$, so

$$\begin{aligned} |\mathbf{u}_h|_h^2 &= \sum_K \|\nabla \mathbf{u}_h\|_{0,K}^2 = \sum_K |K|^2 |f_K|^2 \|\nabla \mathbf{P}_K\|_{0,K}^2 \\ &\leq 1/2 \sum_K |K|\,|f_K|^2 \leq 1/2\|f\|_{0,\Omega}^2 \end{aligned}$$

and we obtain (50).

Next we show (49). By the linearity of p_h, bounds (2) on $\mathcal{A}_K$, and the Cauchy-Schwarz inequality, we get

$$\begin{aligned} |p_h|_h^2 &= \sum_K \int_K |\nabla p_h|^2 dx = \sum_K |K||\nabla p_h|^2 \\ &= \sum_K |K|(\mathcal{A}_K^{-1}\mathbf{u}_K)^T(\mathcal{A}_K^{-1}\mathbf{u}_K) \leq C\sum_K |K|\mathbf{u}_K^T\mathbf{u}_K \\ &\leq C\sum_K \int_K |\mathbf{u}_h(\mathbf{x})|^2 dx = C\|\mathbf{u}_h\|_{0,\Omega}^2, \end{aligned}$$

whence

$$C_1|p_h|_h \leq \|\mathbf{u}_h\|_{0,\Omega}.$$

Again by (14), bound (2) on $\mathcal{A}_K$,

$$\begin{aligned} \|\mathbf{u}_h\|_{0,K} &\leq \|\mathcal{A}_K\nabla p_h\|_{0,K} + |K|\,|f_K|\,\|\mathbf{P}_K\|_{0,K} \\ &= (\int_K \nabla p_h^T \mathcal{A}_K^2 \nabla p_h\, dx)^{1/2} + |K|\,|f_K|\,\|\mathbf{P}_K\|_{0,K} \\ &\leq C\|\nabla p_h\|_{0,K} + |K|\,|f_K|\,\|\mathbf{P}_K\|_{0,K} \\ &= C|p_h|_{h,K} + |K|\,|f_K|\,\|\mathbf{P}_K\|_{0,K} \end{aligned}$$

¿From (13) we have

$$\begin{aligned} \|\mathbf{P}_K\|_{0,K}^2 &= \frac{1}{4|K|^2}\int_K (x-x_B)^2 + (y-y_B)^2 dxdy \\ &\leq \frac{h_K^2}{4|K|}. \end{aligned}$$

By the regularity assumption on the triangulations we know $h_K/|K|^{1/2} \le C$ for all K. Also notice $|f_K| \le \|f\|_{0,K}/|K|^{1/2}$, we obtain

$$\|\mathbf{u}_h\|_{0,K} \le C(|p_h|_{h,K} + |K|^{1/2}\|f\|_{0,K}) \le C(|p_h|_{h,K} + h\|f\|_{0,K}).$$

Summing over K, we get

$$\|\mathbf{u}_h\|_{0,\Omega} \le C(|p_h|_h + h\|f\|_{0,\Omega}).$$

This completes the proof.

Our main result in this section is the following error estimate theorem.

Theorem 5.2 Let the problem data of (1) be smooth enough so that the pressure solution $p \in H^2 \cap H_0^1$ and $\mathbf{u}(\mathbf{x}) = -\mathcal{K}(\mathbf{x})\nabla p(\mathbf{x}) \in H^1(\Omega)^2$. Then there exists a constant C independent of h such that

$$\begin{aligned}
\|p - p_h\|_0 &\le Ch^2(|f|_h + \|f\|_0), \text{ if } f|_K \in H^1(K) \quad \forall K \in \mathcal{T}_h, &(51)\\
|p - p_h|_h &\le Ch\|f\|_0 &(52)\\
\|\mathbf{u} - \mathbf{u}_h\|_0 &\le Ch(\|f\|_0 + \|\mathbf{u}\|_0) &(53)\\
|\mathbf{u} - \mathbf{u}_h|_{H(\mathrm{div};\Omega)} &\le Ch|f|_h \qquad \text{if } f|_K \in H^1(K) \quad \forall K \in \mathcal{T}_h. &(54)
\end{aligned}$$

Also we assume that $\mathcal{K} \in W^{1,\infty}$.

Remark. Note that the condition of the function $f \in H^1(K)$ for all K can be satisfied if we impose it on the coarsest mesh in the context of local refinement.

Proof 7 *i) proof of (51) and (52).* First we show that

$$\|\tilde{p} - p\|_0 \le Ch^2|f|_h$$

where $\tilde{p} \in H_0^1$ is the solution of $a(\tilde{p}, q) = (f_h, q) \quad \forall q \in H_0^1$ and p is the solution of $a(p, q) = (f, q) \quad \forall q \in H_0^1$. Subtracting these two bilinear forms, we get

$$a(p - \tilde{p}, q) = (f - f_h, q) \quad \forall q \in H_0^1.$$

Noticing that $\int_K f - f_h dx = 0$, we have $(f - f_h, q)_K = (f - f_h, q - q_h)_K$ where $q_h|_K = \frac{1}{|K|}\int_K q dx$ is constant on each K. Then by the Cauchy–Schwarz inequality and an interpolation theorem, we have

$$\begin{aligned}
|(f - f_h, q - q_h)| &= |\sum_K (f - f_h, q - q_h)_K| &(55)\\
&\le Ch^2 \sum_K |f|_{1,K}|q|_{1,K}\\
&\le Ch^2|f|_h|q|_1
\end{aligned}$$

Thus

$$|a(p - \tilde{p}, q)| \leq Ch^2|f|_h|q|_1.$$

Taking $q = p - \tilde{p}$ and using the coercivity of $a(\cdot, \cdot)$, we have

$$|p - \tilde{p}|_1^2 \leq Ch^2|f|_h|p - \tilde{p}|_1$$

and hence

$$|p - \tilde{p}|_1 \leq Ch^2|f|_h. \tag{56}$$

By the Poincaré inequality, we get

$$\|p - \tilde{p}\|_0 \leq Ch^2|f|_h. \tag{57}$$

Now note that $p_h \in Y_{h,0}$ is the solution of the nonconforming method

$$a_h(p_h, q_h) = (f_h, q_h) \qquad \forall q_h \in Y_{h,0}$$

associated with the variational problem of finding $\tilde{p} \in H_0^1$ such that

$$a(\tilde{p}, q) = (f_h, q) \qquad \forall q \in H_0^1.$$

(The right hand side contains f_h instead of f!) Hence

$$||p_h - \tilde{p}||_0 + h|\tilde{p} - p_h|_h \leq Ch^2||\tilde{p}||_2, \tag{58}$$

where we have applied (44) with the right side f_h and all the remarks concerning regularity of the solution made there then also apply here.

Now by the triangle inequality, (57), (58), the stability condition and the fact that f_h is the L_2 projection of f, we have

$$\begin{aligned} ||p_h - p||_0 &\leq ||p_h - \tilde{p}||_0 + ||\tilde{p} - p||_0 \\ &\leq Ch^2||\tilde{p}||_2 + Ch^2|f|_h \\ &\leq Ch^2||f_h||_0 + Ch^2|f|_h \\ &\leq Ch^2||f||_0 + Ch^2|f|_h. \end{aligned}$$

This completes the proof of (51).

As for the proof of (52), first note that from (55) we can use the fact that $\|f - f_h\|_{0,K} \leq \|f\|_{0,K}$ and an interpolation theorem to obtain

$$\begin{aligned} |(f - f_h, q - q_h)| &\leq Ch \sum_K \|f\|_{0,K}|q|_{1,K} \\ &\leq Ch\|f\|_0|q|_1 \end{aligned}$$

and consequently instead of (56) we have $|\tilde{p} - p|_1 \leq Ch\|f\|_0$; arguing as before. Now

$$\begin{aligned} |p_h - p|_h &\leq |p_h - \tilde{p}|_h + |\tilde{p} - p|_h \\ &\leq Ch||\tilde{p}||_2 + Ch\|f\|_0. \\ &\leq Ch||f||_0, \end{aligned}$$

ii) *proof of* (53).
From (14) we have

$$\mathbf{u}_h(\mathbf{x}) - \mathbf{u}(\mathbf{x})|_K = -\mathcal{A}_K \nabla p_h + \mathcal{K}(\mathbf{x})\nabla p + |K| f_K \mathbf{P}_K(\mathbf{x}).$$

So

$$\|\mathbf{u}_h - \mathbf{u}\|_{0,K} \leq \|\mathcal{A}_K \nabla p_h - \mathcal{K}(\mathbf{x})\nabla p\|_{0,K} + |K|\,|f_K|\,\|\mathbf{P}_K\|_{0,K},$$

and since $|K|\,|f_K|\ \|\mathbf{P}_K\|_{0,K} \leq Ch\|f\|_{0,K}$ as shown in Lemma 5.2 we have

$$\|\mathbf{u}_h - \mathbf{u}\|_{0,K} \leq \|\mathcal{A}_K \nabla p_h - \mathcal{K}(\mathbf{x})\nabla p\|_{0,K} + Ch\|f\|_{0,K}$$

whereas by the triangle inequality and the interpolation theorem

$$\begin{aligned} \|\mathcal{A}_K \nabla p_h &- \mathcal{K}\nabla p\|_{0,K} \leq \|\mathcal{A}_K \nabla p_h - \mathcal{K}\nabla p_h\|_{0,K} + \|\mathcal{K}\nabla p_h - \mathcal{K}\nabla p\|_{0,K} \\ &= (\nabla p_h^T \int_K (\mathcal{A}_K - \mathcal{K})^2 d\mathbf{x} \nabla p_h)^{1/2} + (\int_K |\mathcal{K}(\nabla p_h - \nabla p)|^2 d\mathbf{x})^{1/2} \\ &\leq Ch|p_h|_{h,K}|\mathcal{K}|_{1,\infty,K} + C|p_h - p|_{E,K} \\ &\leq Ch\|\mathbf{u}_h\|_{0,K} + C|p_h - p|_{E,K} \\ &\leq C\{h\|\mathbf{u} - \mathbf{u}_h\|_{0,K} + h\|\mathbf{u}\|_{0,K} + C|p_h - p|_{E,K}\} \end{aligned}$$

where we have used the stability (49) restricted on K as shown in its proof. Now taking h small enough to move the first term on the right side we have

$$\|\mathbf{u}_h - \mathbf{u}\|_{0,K} \leq C\{|p_h - p|_{E,K} + h\|\mathbf{u}\|_{0,K} + h\|f\|_{0,K}\}.$$

Summing over K and using (52) we have

$$\|\mathbf{u}_h - \mathbf{u}\|_0 \leq Ch(\|f\|_0 + \|\mathbf{u}\|_0).$$

iii) *proof of (54).*
By $(4)_1$, $\nabla \cdot \mathbf{u}_h(\mathbf{x}) = f_K$ and hence

$$\|\nabla \cdot \mathbf{u}_h - \nabla \cdot \mathbf{u}\|_{0,K} = \|f - f_K\|_{0,K} \leq Ch|f|_{1,K} \tag{59}$$

where an interpolation theorem is used. Summing over K, we get

$$|\mathbf{u} - \mathbf{u}_h|_{H(\text{div};\Omega)} \leq \|\nabla \cdot \mathbf{u}_h - \nabla \cdot \mathbf{u}\|_{0,\Omega} \leq Ch|f|_h.$$

6 Numerical Examples

Notice that the the error estimate theorem and the algorithm are valid for both Dirichlet and Neumann problems. We present numerical results for both cases. We partition the unit square $[0, 1] \times [0, 1]$ into squares evenly in both directions with the diagonals running from the upper-left corner of each triangle to its lower-right corner. We use the incomplete LU preconditioned conjugate gradient method to solve all the problems. The integral of f over element K is computed by the midpoint rule using the three edges of the triangle. Our experiments suggest second order approximation in all cases. The discrete norms in which the errors are estimated are as follows.

6.1 Choice of discrete norms

In Thm. 5.2 we predicted first order convergence in the $H(\text{div}; \Omega)$ norm for flux $\mathbf{u}$ and second order convergence in the L_2 norm for the pressure p. We need to choose proper discrete norms to measure the error between true solution and computed solution.

Let (x_i, y_j) be the center of square (i, j) with $x_i = (i - 1/2)h$, $y_j = (j - 1/2)h$, $h = 1/n, i, j = 1, 2, \ldots, n$. Let p_{ij} be the computed pressure at (x_i, y_j). We define

$$pErr\_abs = \|p - p_h\| := \left[\sum_{i,j=1}^{n} h^2 (p(x_i, y_j) - p_{ij})^2 \right]^{\frac{1}{2}},$$

i.e. a discrete L^2–norm of the error $p - p_h$.

Due to (59), the $H(\text{div}; \Omega)$–seminorm of the error in the flux is directly related to $f - f_K$, and so for the flux we will use only an equivalent L_2 discrete norm:

$$uErr\_abs \quad = \quad \|\mathbf{u} - \mathbf{u}_h\| := \left[\sum_K (\int_{\partial K} [(\mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}] \, ds)^2 \right]^{1/2}$$

where the edge integrals are evaluated by the midpoint rule [11]. (Note that the x components are picked out by the vertical edges, the y components by the horizontal edges, etc..)

We also compute the relative error

$$pErr\_rel := pErr\_abs / \|p\|$$

and

$$uErr\_rel := uErr\_abs / \|\mathbf{u}\|.$$

6.2 Dirichlet problems

We consider the following Dirichlet problem

$$\begin{cases} -\nabla \cdot \mathcal{K}\nabla p = f & \text{in } \Omega, \\ p = 0 & \text{on } \partial\Omega. \end{cases} \tag{60}$$

In examples 1–4, the true pressure is $p = (x^2 - x)(y^2 - y)$ on the unit square.
Example 1. The mobility tensor $\mathcal{K} = diag(1 + 10x^2 + y^2, 1 + x^2 + 10y^2)$.

Table 1: Error behavior for Dirichlet problem

Example 1.	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$	order
pErr_abs	8.7748e-5	2.2318e-5	5.6041e-6	1.4026e-6	≈ 2
pErr_rel	0.0026	6.6952e-4	1.6812e-4	4.2077e-5	≈ 2
uErr_abs	0.0113	0.0029	7.1931e-4	1.8029e-4	≈ 2
uErr_rel	0.0086	0.0021	5.1929e-4	1.2905e-4	≈ 2
Length of P	736	3,008	12,160	48,896	
Length of U	736	3,008	12,160	48,896	

Example 2. The mobility tensor $\mathcal{K} = diag(10^4, 1)$.

Table 2: Error behavior for Dirichlet problem

Example 2.	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$	order
pErr_abs	9.1815e-5	2.3286e-5	5.8558e-6	1.4657e-6	≈ 2
pErr_rel	0.0028	6.9859e-4	1.7567e-4	4.3971e-5	≈ 2
uErr_abs	8.0759	2.0653	0.5218	0.1311	≈ 2
uErr_rel	0.0057	0.0014	3.5417e-4	8.8436e-5	≈ 2
Length of P	736	3,008	12,160	48,896	
Length of U	736	3,008	12,160	48,896	

Example 3. In this example, we study the effect of discontinuous mobility matrix $\mathcal{K}$. Let
$\mathcal{K} = \begin{bmatrix} 10000 & 0 \\ 0 & 1 \end{bmatrix}$ on the left half unit square, and $\mathcal{K} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ on the right half unit square. We find the approximation is as good as continuous cases.
Example 4. Let $\mathcal{K} = \begin{bmatrix} 1 + 10x^2 + y^2 & 1/2 + x^2 + y^2 \\ 1/2 + x^2 + y^2 & 1 + x^2 + 10y^2 \end{bmatrix}$.

Table 3: Error behavior for Dirichlet problem

Example 3.	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$	order
pErr_abs	1.6425e-4	4.1581e-5	1.0439e-5	2.6128e-6	≈ 2
pErr_rel	0.0049	0.0012	3.1317e-4	7.8382e-5	≈ 2
uErr_abs	6.2337	1.5882	0.4005	0.1005	≈ 2
uErr_rel	0.0062	0.0015	3.8441e-4	9.5877e-5	≈ 2
Length of P	736	3,008	12,160	48,896	
Length of U	736	3,008	12,160	48,896	

Table 4: Error behavior for Dirichlet problem

Example 4.	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$	order
pErr_abs	1.4595e-4	3.7458e-5	9.4305e-6	2.3620e-6	≈ 2
pErr_rel	0.0044	0.0011	2.8292e-4	7.0861e-5	≈ 2
uErr_abs	0.0168	0.0043	0.0011	2.7533e-4	≈ 2
uErr_rel	0.0114	0.0028	7.0927e-4	1.7784e-4	≈ 2
Length of P	736	3,008	12,160	48,896	
Length of U	736	3,008	12,160	48,896	

6.3 Neumann problems

$$\begin{cases} -\nabla \cdot \mathcal{K} \nabla p &= f \text{ in } \Omega, \\ \mathcal{K} \nabla p \cdot \mathbf{n} &= 0 \text{ on } \partial\Omega. \end{cases} \tag{61}$$

Example 5. The true pressure is $p = \cos(2\pi x)\cos(2\pi y)$ and $\mathcal{K} = diag(\cos(2\pi y) + 2, \cos(2\pi x) + 2)$.

Table 5: Error behavior for Neumann problem

Example 5.	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$	order
pErr_abs	0.0110	0.0028	6.9570e-4	1.7399e-4	≈ 2
pErr_rel	0.0221	0.0056	0.0014	3.4798e-4	≈ 2
uErr_abs	0.1383	0.0353	0.0089	0.0022	≈ 2
uErr_rel	0.0101	0.0026	6.4814e-4	1.6226e-4	≈ 2
Length of P	800	3,136	12,416	49,408	
Length of U	800	3,136	12,416	49,408	

Example 6. The true (oscillatory) pressure is $p = \cos(2\pi x)\cos(10\pi y)$ and $\mathcal{K} = I$.

Table 6: Error behavior for Neumann problem

Example 6.	$h = 1/16$	$h = 1/32$	$h = 1/64$	$h = 1/128$	order
pErr_abs	0.0441	0.0130	0.0034	8.4912e-4	≈ 2
pErr_rel	0.0882	0.0260	0.0067	0.0017	≈ 2
uErr_abs	4.1789	1.1274	0.2882	0.0725	≈ 2
uErr_rel	0.1845	0.0498	0.0127	0.0032	≈ 2
Length of P	800	3,136	12,416	49,408	
Length of U	800	3,136	12,416	49,408	

References

[1] I. Aavatsmark, T. Barkve, O. Bϕe, and T. Mannseth, *Discretization on unstructured grids for inhomogeneous anisotropic media, part I: derivation of the methods,* SIAM J. Sci. Comput. Vol. 19, No 5, 1700-1716 (1998).

[2] D. N. Arnold and F. Brezzi,*Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates,* RAIRO Model. Math. Anal. Numer. 19, 7-32 (1985).

[3] R. E. Bank and D. J. Rose, *Some error estimates for the box method* , SIAM J. Numer. Anal. 24, 4, 777-787 (1987).

[4] S. C. Brenner, *An optimal-order multigrid method for $P1$ nonconforming finite elements*, Math. Comp. 52, No. 165, 1-15 (1989).

[5] S. C. Brenner and L. Ridgway Scott,*The mathematical theory of finite element methods*, Spring-Verlag, (1994).

[6] Z. Cai, J. E. Jones, S. F. McCormick and T. F. Russell, *Control volume mixed finite element methods*, Computational Geosciences 1, 289-315 (1997).

[7] M. Crouzeix and P. A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations I,* R.A.I.R.O. 33-76 (1973).

[8] S. H. Chou, *Analysis and convergence of a covolume method for the generalized Stokes problem,* Math. Comp. 66, 85-104 (1997).

[9] S. H. Chou and D. Y. Kwak, *Mixed covolume methods on rectangular grids for elliptic problems,* SIAM J. Numer. Anal., Vol 37, No 3, 758-771, (2000).

[10] S. H. Chou, D. Y. Kwak and P.S. Vassilevski, *Mixed covolume methods for elliptic problems on triangular grids,* SIAM J. Numer. Anal., Vol. 35, No. 5, 1850-1861, (1998).

[11] S. H. Chou and P. S. Vassilevski, *A General mixed covolume framework for constructing conservative schemes for elliptic problems,* Math. Comp., Vol. 68, No. 227, 991-1011, (1999).

[12] S. H. Chou and D. Y. Kwak, *A covolume method based on rotated bilinears for the generalized Stokes problem,* SIAM J. Numer. Anal., Vol 35, No. 2, 497-507 (1998).

[13] S. H. Chou and S. Tang, *The P1 nonconforming FEM with accurate flux for elliptic problems without using mixed methods is a finite volume box method*, BGSU report, (1999).

[14] S. H. Chou and S. Tang, *Conservative P1 conforming and nonconfirming Galerkin FEMs: effective flux evaluation via a non-mixed method approach,* SIAM J. Numer. Anal., Vol. 38, No. 2, 660-680 (2000).

[15] B. Courbet and J. P. Croisille, *Finite volume box schemes on triangular meshes*, RAIRO Model. Math. Anal. Numer. Vol.32, No. 5, 631-649 (1998)

[16] H. Jianguo and X. Shitong, *On the finite volume element method for general self-adjoint elliptic problems*, SIAM J. Numer. Anal., 35, 1762-1774, (1998).

[17] I. D. Mishev, *Finite volume methods on Voronoi meshes*, Numer. Methods Partial Differential Equations, 14, 2, 193-212 (1998).

[18] R. A. Nicolaides, T. A. Porsching and C. A. Hall, *Covolume methods in computational fluid dynamics*, in Computational Fluid Dynamics Review, M. Hafez and K. Oshma ed., John Wiley and Sons, 279-299, (1995).

[19] P. A. Raviart and J. M. Thomas, *A mixed finite element methods for second order elliptic problems*, in Mathematical aspects of the finite element method, Lecture Notes in Mathematics, 606, 292–315, Spring-Verlag, (1977).

[20] T. F. Russell and M. Wheeler, *Finite element and finite difference methods for continuous flows in porous media,* The Mathematics of Reservoir Simulation, R. E. Ewing, ed., Frontiers in Applied Mathematics 1, Society for Industrial and Applied Mathematics, Philadelphia, (1984).

[21] J. Shen, *A block finite difference scheme for second-order elliptic problems with discontinuous coefficient,* SIAM J. Numer. Anal, 33, 688-706 (1996).

[22] P. S. Vassilevski, S. I. Petrova and R. D. Lazarov, *Finite difference schemes on triangular cell-centered grids with local refinement*, SIAM J. Sci. Stat. Comput. 13, 6, 1287-1313, (1992).

Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, Ohio, U.S.A
email: chou@zeus.bgsu.edu
http://www-math.bgsu.edu/ chou/.

Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, Ohio, U.S.A
email: tang@bgnet.bgsu.edu