J Sci Comput (2012) 52:563–587
DOI 10.1007/s10915-011-9558-3

Mixed Element Method for Two-Dimensional Darcy-Forchheimer Model

Hao Pan · Hongxing Rui

Received: 7 August 2010 / Revised: 8 May 2011 / Accepted: 8 November 2011 /
Published online: 23 November 2011

Abstract A mixed element method is introduced to solve Darcy-Forchheimer equation, in which the velocity and pressure are approximated by mixed element such as Raviart-Thomas, Brezzi-Douglas-Marini element. We establish the existence and uniqueness of the problem. Error estimates are presented based on the monotonicity owned by the Forchheimer term. An iterative scheme is given for practical computation. The numerical experiments using the lowest order Raviart-Thomas (RT_0) mixed element show that the convergence rates of our method are in agree with the theoretical analysis.

Keywords Darcy-Forchheimer's model · Mixed element · Monotone operator · Numerical analysis

1 Introduction

Darcy flow in porous media is of great interest in many fields such as oil recovery and groundwater pollution contamination. Darcy's law,

$$\frac{\mu}{\rho}K^{-1}\boldsymbol{u} + \nabla p = \boldsymbol{g},$$

describes the linear relationship between the velocity of creep flow and the gradient of pressure, which can be derived by a linear simplification of momentum theorem. The relationship is valid by experiment under the condition that the creeping velocity is low and

H. Pan · H. Rui (✉)
School of Mathematics, Shandong University, Jinan 250100, China
e-mail: hxrui@sdu.edu.cn

the porosity and permeability is small enough by Darcy in 1856 [1]. A theoretical derivation of Darcy's law can be found in [18, 24].

When the velocity is higher and the porosity is nonuniform, a nonlinear relationship between velocity and the pressure gradient is developed, suggested by Forchheimer in 1901 [1], by adding a second order term to reach a modified equation. The Forchheimer equation is as follows,

$$\frac{\mu}{\rho}K^{-1}\boldsymbol{u}+\frac{\beta}{\rho}|\boldsymbol{u}|\boldsymbol{u}+\nabla p=\boldsymbol{g}.$$

A theoretical derivation of Forchheimer's law can be found in [21].

Forchheimer's law mainly describes the inertial effects and high speed flow. The most important feature of Forchheimer's law is that it combines the monotonicity of the nonlinear term and the non-degenerate of the Darcy's part. This will be useful in the proofs of existence, uniqueness and error estimates.

Mixed element methods for Darcy equation is the same as the mixed element methods for second order elliptic equation, except the effect of gravitational term. These work can be found in many references such as [3, 23].

A mixed element for Forchheimer equation (or called Darcy-Forchheimer equation sometimes) was introduced by Girault and Wheeler [13]. They proved the existence and uniqueness of a weak solution for the Forchheimer equation. While, their mixed approximation is called primal [20], and they approximated velocity by piecewise constants and pressure by Crouzeix-Raviart element. They also proposed an alternating directions iterative method to solve the system of nonlinear equations obtained by finite element discretizaton. The convergence of both the iterative algorithm and the mixed element scheme are presented, and the error estimate of the mixed element scheme is demonstrated too. Sooner after that, the numerical experiments are carried out in [15], in which the convergence of the approximation and the iterative algorithm are presented.

The objective of this paper is to propose a mixed element approximation, which is different from the scheme in [13]. The mixed formulation used here is different from [13]. The mixed elements used here are the usually used elements such as the Raviart-Thomas mixed element, Brezzi-Douglas-Marini mixed element and so on. We demonstrate the existence and uniqueness of the weak solution, and give the error estimate based on the monotonicity of the Forchheimer term. The second goal of this paper is to carry out numerical tests of the methods proposed in this paper. On one hand, we testify the convergence of the finite element approximation. On the other hand, we look into the nonlinearity of Forchheimer model compared to the Darcy model.

The paper is organized as follows. In Sect. 2 we give some notations and presents the mixed weak formulation and its mixed element approximation. In Sect. 3 we establish the existence and uniqueness of the mixed weak formulation and its mixed element approximation. In Sect. 4 we present the error estimates using the monotonicity owned by the Forchheimer term. In Sect. 5 a simple linear iterative scheme for practical implementation is given. In Sect. 6 some numerical experiment using the lowest order Raviart-Thomas finite element are carried out; we test several examples using triangle and rectangle element in two-dimension. The numerical result shows that the convergence of our method is in agree with the theoretical analysis and shows the different behaviors between Darcy equation and Forchheimer equation.

2 Mixed Weak Formulation and Mixed Element Approximation

In this section we will present the mixed weak formulation and mixed element approximation of the Forchheimer equation. We consider the following model problem

$$\begin{cases} \text{(i)} & \frac{\mu}{\rho}K^{-1}\boldsymbol{u} + \frac{\beta}{\rho}|\boldsymbol{u}|\boldsymbol{u} + \nabla p = \boldsymbol{g}, & \boldsymbol{x} \in \Omega, \\ \text{(ii)} & \nabla \cdot \boldsymbol{u} = f, & \boldsymbol{x} \in \Omega, \\ \text{(iii)} & p = f_D, & \boldsymbol{x} \in \partial\Omega. \end{cases} \tag{1}$$

Here p represents the pressure while $\boldsymbol{u}$ represents the velocity. Ω is a bounded subset of R^d $(d = 2, 3)$ with Lipschitz continuous boundary Γ. For example, Ω is a polygon in two-dimension or polyhedron in three-dimension. $\boldsymbol{n}$ is the unit exterior normal vector to the boundary of Ω. $|\cdot|$ denotes the Euclidean norm, $|\boldsymbol{u}|^2 = \boldsymbol{u} \cdot \boldsymbol{u}$. ρ, μ and β are scalar functions that represent the density of the fluid, its viscosity and its dynamic viscosity, respectively. β is also referred as Forchheimer number when it is a scalar positive constant. $K = \begin{pmatrix} k_1 & \\ & k_2 \end{pmatrix}$ is the permeability tensor function.

We suppose that $\mu(\boldsymbol{x}), \rho(\boldsymbol{x}), \beta(\boldsymbol{x}) \in L^\infty(\Omega)$ and $K(\boldsymbol{x}) \in (L^\infty(\Omega))^{2\times 2}$. We suppose further that all the coefficients above satisfying the following conditions.

$$\begin{aligned} 0 < \mu_{min} \le \mu(\boldsymbol{x}) \le \mu_{max}, \\ 0 < \rho_{min} \le \rho(\boldsymbol{x}) \le \rho_{max}, \\ 0 < \beta_{min} \le \beta(\boldsymbol{x}) \le \beta_{max}, \end{aligned}$$

$K(\boldsymbol{x})$ is uniformly positively defined and bounded, i.e.

$$0 < K_{min}\boldsymbol{x} \cdot \boldsymbol{x} \le (K(\boldsymbol{x})\boldsymbol{x}) \cdot \boldsymbol{x} \le K_{max}\boldsymbol{x} \cdot \boldsymbol{x}.$$

$\boldsymbol{g}(\boldsymbol{x}) = \nabla Z(\boldsymbol{x}) \in (L^2(\Omega))^d$, a vector function, is the gradient of the depth function $Z(\boldsymbol{x}) \in H^1(\Omega)$. $f(\boldsymbol{x}) \in L^2(\Omega)$, a scalar function, represents the source and sink of the systems. $f_D(\boldsymbol{x}) \in L^2(\partial\Omega)$, a scalar function, represents the Dirichlet boundary condition.

Remark The boundary condition of problem (1) can be replaced by Neumann boundary condition, i.e.

$$\boldsymbol{u} \cdot \boldsymbol{n} = f_N, \qquad \boldsymbol{x} \in \partial\Omega. \tag{2}$$

In this case, a compatibility condition should be given as bellow:

$$\int_\Omega f \, \mathrm{d}\boldsymbol{x} = \int_{\partial\Omega} f_N \, \mathrm{d}s. \tag{3}$$

In addition, the Dirichlet boundary condition appears as natural condition and the Neumann boundary condition appears as essential condition here. It is contrary to the primal mixed element [13], under which the Dirichlet boundary condition seems as a essential condition and the Neumann boundary condition seems as a natural condition [20].

Define the function spaces X, M and their norms as follows:

$$\begin{aligned} &X = \{\boldsymbol{u} \in L^3(\Omega)^d; \nabla \cdot \boldsymbol{u} \in L^2(\Omega)\}, \qquad M = L^2(\Omega), \\ &\|\boldsymbol{u}\|_X = \|\boldsymbol{u}\|_{0,3,\Omega} + \|\nabla \cdot \boldsymbol{u}\|_{0,2,\Omega}, \qquad \|p\|_M = \|p\|_{0,2,\Omega}. \end{aligned}$$

Table 1 RT type mixed elements

dim	Element	$X_h(T)$	$M_h(T)$
2D	Triangle	$RT_k(T) = P_k(T)^2 \oplus \boldsymbol{x} P_k(T)$	$P_k(T)$
2D	Rectangle	$RT_{[k]}(T) = Q_{k+1,k}(T) \oplus Q_{k,k+1}(T)$	$Q_{k,k}(T)$
3D	Tetrahedron	$RTN_k(T) = P_k(T)^3 \oplus \boldsymbol{x} P_k(T)$	$P_k(T)$
3D	Cuboid	$RTN_{[k]}(T) = Q_{k+1,k,k}(T) \oplus Q_{k,k+1,k}(T) \oplus Q_{k,k,k+1}(T)$	$Q_{k,k,k}(T)$

Notes: $P_k(T)$ is the polynomial up to k order in d dimension used in triangle or tetrahedron, while $Q_{k,l}(T)$ or $Q_{k,l,m}(T)$ are the polynomial up to k, l, m order in each dimension ($d = 2$ or $d = 3$) used in rectangle or cuboid

The subscript Ω will always be omitted unless necessary to avoid ambiguity.

Using the integration by parts, we get

$$\int_\Omega \left(\frac{\mu}{\rho} K^{-1}\boldsymbol{u} + \frac{\beta}{\rho}|\boldsymbol{u}|\boldsymbol{u}\right)\cdot \boldsymbol{v}\,\mathrm{d}\boldsymbol{x} - \int_\Omega p\nabla\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} + \int_{\partial\Omega} p\boldsymbol{v}\cdot\boldsymbol{n}\,\mathrm{d}s = \int_\Omega \boldsymbol{g}\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}, \quad \forall \boldsymbol{v}\in X.$$

Applying the Dirichlet boundary condition, we list the mixed weak form as follows

$$(P)\begin{cases} \text{Find } \boldsymbol{u}\in X,\ p\in M \text{ such that,} \\ \text{(i)} \quad \int_\Omega \left(\frac{\mu}{\rho}K^{-1}\boldsymbol{u} + \frac{\beta}{\rho}|\boldsymbol{u}|\boldsymbol{u}\right)\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} - \int_\Omega p\nabla\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \\ \qquad\quad = -\int_{\partial\Omega} f_D\boldsymbol{v}\cdot\boldsymbol{n}\,\mathrm{d}s + \int_\Omega \boldsymbol{g}\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}, \quad \forall\boldsymbol{v}\in X, \\ \text{(ii)} \qquad -\int_\Omega w\nabla\cdot\boldsymbol{u}\,\mathrm{d}\boldsymbol{x} = -\int_\Omega wf\,\mathrm{d}\boldsymbol{x}, \quad \forall w\in M. \end{cases} \tag{4}$$

Let $\mathcal{T}_h$ be a quasi-regular polygonalization of Ω (by triangles, rectangles, tetrahedron or possibly hexahedron), with h being the maximum diameter of the elements of the polygonalization. Let $X_h \times M_h \subset X \times M$ be a conforming (or compatible) mixed element space with index k and discretization parameter h. $X_h \times M_h$ is an approximation to $X \times M$.

There are many conforming (or compatible) mixed element function spaces such as Raviart-Thomas [19], Brezzi-Douglas-Marini [6] in 2D triangle and rectangle; Raviart-Thomas-Nedelec [16, 17], Brezzi-Douglas-Duran-Fortin [4] in 3D tetrahedron and cuboid; Brezzi-Douglas-Fortin-Marini [5] in 2D rectangle or 3D cuboid; they are summarized in many books [7, 8, 20]. Some RT type mixed element are listed in Table 1.

Replacing the original velocity and pressure by their approximate, we get the mixed element approximation problem,

$$(P_h)\begin{cases} \text{Find } \boldsymbol{u}_h\in X_h,\ p_h\in M_h \text{ such that,} \\ \text{(i)} \quad \int_\Omega \left(\frac{\mu}{\rho}K^{-1}\boldsymbol{u}_h + \frac{\beta}{\rho}|\boldsymbol{u}_h|\boldsymbol{u}_h\right)\cdot\boldsymbol{v}_h\,\mathrm{d}\boldsymbol{x} - \int_\Omega p_h\nabla\cdot\boldsymbol{v}_h\,\mathrm{d}\boldsymbol{x} \\ \qquad\quad = -\int_{\partial\Omega} f_D\boldsymbol{v}_h\cdot\boldsymbol{n}\,\mathrm{d}s + \int_\Omega \boldsymbol{g}\cdot\boldsymbol{v}_h\,\mathrm{d}\boldsymbol{x}, \quad \forall\boldsymbol{v}_h\in X_h, \\ \text{(ii)} \qquad -\int_\Omega w_h\nabla\cdot\boldsymbol{u}_h\,\mathrm{d}\boldsymbol{x} = -\int_\Omega w_h f\,\mathrm{d}\boldsymbol{x}, \quad \forall w_h\in M_h. \end{cases} \tag{5}$$

Now we define the interpolation or projection below. For velocity, let $\Pi_h : X \to X_h$ be the Raviart-Thomas projection [19] or Brezzi-Douglas-Marini projection [6], which satisfies

$$\int_\Omega w_h \nabla \cdot (\Pi_h \boldsymbol{v} - \boldsymbol{v})\,\mathrm{d}\boldsymbol{x} = 0, \quad \boldsymbol{v} \in X,\ w_h \in M_h; \tag{6}$$

$$\|\Pi_h \boldsymbol{v} - \boldsymbol{v}\|_{0,q} \le C\|\boldsymbol{v}\|_{s,q} h^s, \quad \frac{1}{q} < s \le k+1,\ \forall \boldsymbol{v} \in X \cap W^{s,q}(\Omega)^d; \tag{7}$$

$$\|\nabla \cdot (\Pi_h \boldsymbol{v} - \boldsymbol{v})\|_{0,2} \le C\|\nabla \cdot \boldsymbol{v}\|_{s,2} h^s, \quad 0 \le s \le k+1,\ \forall \boldsymbol{v} \in X \cap H^s(\mathrm{div}, \Omega). \tag{8}$$

For pressure, let $P_h : M \to M_h$ be the orthogonal L^2 projection, which satisfies

$$\int_\Omega (P_h w - w)\chi_h\,\mathrm{d}\boldsymbol{x} = 0, \quad w \in M,\ \forall \chi_h \in M_h; \tag{9}$$

$$\|P_h w - w\|_{0,q} \le C\|w\|_{s,q} h^s, \quad 0 \le s \le k+1,\ \forall w \in M \cap W^{s,q}(\Omega); \tag{10}$$

$$\|P_h w - w\|_{-r,2} \le C\|w\|_s h^{r+s,2}, \quad 0 \le r, s \le k+1,\ \forall w \in H^s(\Omega). \tag{11}$$

The two projections above preserve the relation

$$\int_\Omega (P_h w - w)\nabla \cdot \boldsymbol{v}_h\,\mathrm{d}\boldsymbol{x} = 0, \quad w \in M,\ \boldsymbol{v}_h \in X_h, \tag{12}$$

i.e. $\mathrm{div}\, X_h = M_h$. The operator form,

$$\Pi_h \times P_h : X \times M \to X_h \times M_h, \tag{13}$$

has the commuting property: $\mathrm{div} \circ \Pi_h = P_h$, i.e.

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{div}} & M \\ {\scriptstyle \Pi_h}\downarrow & & \downarrow{\scriptstyle P_h} \\ X_h & \xrightarrow{\mathrm{div}} & M_h. \end{array} \tag{14}$$

For simplicity we rewrite the variational form (4) and (5) using inner product format. Define some symbols first

$$D^{-1}(\boldsymbol{u}) = \frac{\mu}{\rho K} + \frac{\beta}{\rho}|\boldsymbol{u}|, \quad \text{if } K \text{ is scalar;}$$

$$D^{-1}(\boldsymbol{u}) = \frac{\mu}{\rho} K^{-1} + \frac{\beta}{\rho}|\boldsymbol{u}| I, \quad \text{if } K \text{ is tensor;}$$

$$A(\boldsymbol{u}) = D^{-1}(\boldsymbol{u})\boldsymbol{u}.$$

We rewrite the weak form (4) and the mixed element approximation (5) as follows.

$$(\tilde{P}) \begin{cases} \text{Find } \boldsymbol{u} \in X,\ p \in M \text{ such that,} \\ \text{(i)} \quad (A(\boldsymbol{u}), \boldsymbol{v}) - (p, \nabla \cdot \boldsymbol{v}) = (\boldsymbol{g}, \boldsymbol{v}) - (f_D, \boldsymbol{v} \cdot \boldsymbol{n}), \quad \forall \boldsymbol{v} \in X, \\ \text{(ii)} \quad -(w, \nabla \cdot \boldsymbol{u}) = -(w, f), \quad \forall w \in M; \end{cases} \tag{15}$$

$$(\tilde{P}_h)\begin{cases}\text{Find } \boldsymbol{u}_h \in X_h,\ p_h \in M_h \text{ such that,}\\ \text{(i)}\quad (A(\boldsymbol{u}_h), \boldsymbol{v}_h) - (p_h, \nabla\cdot \boldsymbol{v}_h) = (\boldsymbol{g}, \boldsymbol{v}_h) - (f_D, \boldsymbol{v}_h\cdot \boldsymbol{n}), \quad \forall \boldsymbol{v}_h \in X_h,\\ \text{(ii)}\qquad\qquad -(w_h, \nabla\cdot \boldsymbol{u}_h) = -(w_h, f), \quad \forall w_h \in M_h.\end{cases} \tag{16}$$

3 Existence and Uniqueness

For simplicity, we just consider the Dirichlet boundary condition with $f_D = 0$ on $\partial\Omega$. We assume that the permeability is a scalar. Then

$$\left(\frac{\mu}{\rho K} + \frac{\beta}{\rho}|\boldsymbol{u}|\right)|\boldsymbol{u}| = |-\nabla p + \boldsymbol{g}|,$$

$$\frac{\beta}{\rho}|\boldsymbol{u}|^2 + \frac{\mu}{\rho K}|\boldsymbol{u}| - |\nabla p - \boldsymbol{g}| = 0.$$

Dropping the negative term, we get $|\boldsymbol{u}|$,

$$|\boldsymbol{u}| = \frac{-\frac{\mu}{\rho K} + \sqrt{(\frac{\mu}{\rho K})^2 + 4\frac{\beta}{\rho}|\nabla p - \boldsymbol{g}|}}{2\frac{\beta}{\rho}}.$$

The flux $\boldsymbol{u}$ can then be written as follows

$$\boldsymbol{u} = -\frac{\nabla p - \boldsymbol{g}}{\frac{\mu}{\rho K} + \frac{\beta}{\rho}|\boldsymbol{u}|} = -\frac{2(\nabla p - \boldsymbol{g})}{\frac{\mu}{\rho K} + \sqrt{(\frac{\mu}{\rho K})^2 + 4\frac{\beta}{\rho}|\nabla p - \boldsymbol{g}|}},$$

then Darcy-Forchheimer Equation can be written as follows

$$-\nabla\cdot\left(\frac{2(\nabla p - \boldsymbol{g})}{\frac{\mu}{\rho K} + \sqrt{(\frac{\mu}{\rho K})^2 + 4\frac{\beta}{\rho}|\nabla p - \boldsymbol{g}|}}\right) = f. \tag{17}$$

The Darcy-Forchheimer equation is a monotone nonlinear problem, and under mild regularity assumptions on the data, it has been demonstrated that a unique weak solution exists. Some result are listed below. To prove the existence and uniqueness of the problem (4) or (15), we shall use the following abstract result in [12].

Lemma 3.1 *Let $(X, \|\cdot\|_X)$ and $(M, \|\cdot\|_M)$ be two reflexive Banach spaces. Let $(X', \|\cdot\|_{X'})$ and $(M', \|\cdot\|_{M'})$ be their corresponding dual spaces. Let $B : X \to M'$ be a linear continuous operator and $B' : M \to X'$ the dual operator of B. Let $V = Ker(B)$ be the kernel of B; denotes by $V^0 \subset X'$ the polar set of V, $V^0 = \{x' \in X' | \langle x', v\rangle = 0, \forall v \in X\}$ and $\dot{B} : (X/V) \to M'$ the quotient operator associated with B. The three following properties are equivalent*:

(i) *There exists a constant $\gamma > 0$ such that*

$$\inf_{w\in M}\sup_{v\in X}\frac{\langle Bv, w\rangle}{\|w\|_M\|v\|_X} \geq \gamma,$$

(ii) *B' is an isomorphism from M onto V^0 and*

$$\|B'w\|_{X'} \geq \gamma \|w\|_M, \quad \forall w \in M,$$

(iii) *$\dot{B}$ is an isomorphism form (X/V) onto M' and*

$$\|\dot{M}\dot{v}\|_{M'} \geq \gamma \|\dot{v}\|_{(X/V)}, \quad \forall \dot{v} \in (X/V).$$

The following two lemmas are similar to lemmas in [10, 11], in which the monotone type problem is also studied.

Lemma 3.2 (Inf-sup condition of the continuous problem) *There exists a positive constant γ such that*

$$\inf_{w\in M} \sup_{\boldsymbol{v}\in X} \frac{(w, \nabla \cdot \boldsymbol{v})}{\|w\|_M \|\boldsymbol{v}\|_X} \geq \gamma. \tag{18}$$

Proof For $w \in L^2(\Omega)$, let p satisfy

$$\begin{cases} -\nabla \cdot \left(\dfrac{2(\nabla p - \boldsymbol{g})}{\frac{\mu}{\rho K} + \sqrt{(\frac{\mu}{\rho K})^2 + 4\frac{\beta}{\rho}|\nabla p - \boldsymbol{g}|}} \right) = w, & \text{in } \Omega, \\ p = 0, & \text{on } \partial\Omega. \end{cases}$$

For the problem is nondegenerate and monotone elliptic problem, similar to [9–11] it is easy to get that there exists a constant C such that

$$\|p - Z\|_{0,\frac{3}{2}} + \|\nabla p - \boldsymbol{g}\|_{0,\frac{3}{2}} \leq C\|w\|_{0,2}^2.$$

Therefore, $\boldsymbol{v} \equiv -\frac{2(\nabla p - \boldsymbol{g})}{\frac{\mu}{\rho K} + \sqrt{(\frac{\mu}{\rho K})^2 + 4\frac{\beta}{\rho}|\nabla p - \boldsymbol{g}|}}$ is in the space X and there exists a constant C such that

$$\begin{aligned} \|\boldsymbol{v}\|_X &= \|\boldsymbol{v}\|_{0,3} + \|\nabla \cdot \boldsymbol{v}\|_{0,2} \\ &\leq \left(\int_\Omega \frac{(|2(\nabla p - \boldsymbol{g})|)^3}{(\sqrt{4\frac{\beta}{\rho}|\nabla p - \boldsymbol{g}|})^3} \, \mathrm{d}\boldsymbol{x} \right)^{\frac{1}{3}} + \|\nabla \cdot \boldsymbol{v}\|_{0,2} \\ &\leq C \left(\int_\Omega |\nabla p - \boldsymbol{g}|^{\frac{3}{2}} \, \mathrm{d}\boldsymbol{x} \right)^{\frac{1}{3}} + \|\nabla \cdot \boldsymbol{v}\|_{0,2} \\ &\leq C \|\nabla p - \boldsymbol{g}\|_{0,\frac{3}{2}}^{\frac{1}{2}} + \|w\|_{0,2} \\ &\leq C\|w\|_{0,2}. \end{aligned}$$

Hence $\|\boldsymbol{v}\|_X \leq C\|w\|_{0,2}$, and $(w, \nabla \cdot \boldsymbol{v}) = \|w\|_{0,2}^2$. It is easy to get

$$\frac{(w, \nabla \cdot \boldsymbol{v})}{\|w\|_M \|\boldsymbol{v}\|_X} \geq \frac{\|w\|_{0,2}^2}{C\|w\|_{0,2}^2}.$$

Thus, continuous inf-sup condition (18) follows immediately. □

Theorem 3.3 (Existence and uniqueness of the continuous problem) *The problem* (P) *has a unique solution* $(\boldsymbol{u}, p) \in X \times M$, *and there exists a constant* $C > 0$ *such that*

$$\|\boldsymbol{u}\|_{0,2}^2 + \|\boldsymbol{u}\|_{0,3}^3 + \|\nabla \cdot \boldsymbol{u}\|_{0,2}^2 \leq C(\|f\|_{0,2}^2 + \|f\|_{0,2}^3 + \|\boldsymbol{g}\|_{0,2}^2), \tag{19}$$

$$\|p\|_{0,2} \leq C(\|f\|_{0,2} + \|f\|_{0,2}^2 + \|\boldsymbol{g}\|_{0,2} + \|\boldsymbol{g}\|_{0,2}^2). \tag{20}$$

Proof Let

$$K(f) = \{\boldsymbol{u} \in X;\ (w, \nabla \cdot \boldsymbol{u}) - (w, f) = 0,\ \forall w \in M\}$$

and let $B = \nabla \cdot : X \to M$. Using the inf-sup condition and applying Lemma 3.1 to the operator B, we get the existence of $\boldsymbol{u}_0 \in K(f)$ such that $\|\boldsymbol{u}_0\|_X \leq \frac{1}{\gamma}\|f\|_{0,2}$. Therefore, problem may be written in the following way :

$$\begin{cases} \text{Find } \boldsymbol{u}_1 = (\boldsymbol{u} - \boldsymbol{u}_0) \in X,\ p \in M \text{ such that} \\ \quad (A(\boldsymbol{u}), \boldsymbol{v}) - (p, \nabla \cdot \boldsymbol{v}) = (\boldsymbol{g}, \boldsymbol{v}), \quad \forall \boldsymbol{v} \in X, \\ \qquad\qquad -(w, \nabla \cdot \boldsymbol{u}) = -(w, f), \quad \forall w \in M \end{cases} \tag{21}$$

or

$$\begin{cases} \text{Find } \boldsymbol{u}_1 \in K(0) \text{ such that} \\ (A(\boldsymbol{u}_1 + \boldsymbol{u}_0), \boldsymbol{v}) = (\boldsymbol{g}, \boldsymbol{v}), \quad \forall \boldsymbol{v} \in K(0). \end{cases} \tag{22}$$

Let J be the functional defined by

$$J(\boldsymbol{v}) = \frac{1}{2}\frac{\mu}{\rho}\int_\Omega (K^{-1}\boldsymbol{v}) \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} + \frac{1}{3}\frac{\beta}{\rho}\int_\Omega |\boldsymbol{v}|^3\, \mathrm{d}\boldsymbol{x} - \int_\Omega \boldsymbol{g} \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{x}.$$

Then problem is equivalent to

$$\begin{cases} \text{Find } \boldsymbol{u}_1 \in K(0) \text{ such that} \\ J(\boldsymbol{u}_1 + \boldsymbol{u}_0) = \inf\limits_{\boldsymbol{v} \in K(0)} J(\boldsymbol{v} + \boldsymbol{u}_0). \end{cases} \tag{23}$$

From the properties of J [13], one can deduce that problem has a unique solution $\boldsymbol{u}_1$. Furthermore, there exists a constant C and a arbitrarily small constant ε

$$\begin{aligned} \|\boldsymbol{u}\|_{0,2}^2 + \|\boldsymbol{u}\|_{0,3}^3 - C\|\boldsymbol{g}\|_{0,2}^2 - \varepsilon\|\boldsymbol{u}\|_{0,2}^2 &\leq CJ(\boldsymbol{u}_1 + \boldsymbol{u}_0) \leq CJ(\boldsymbol{u}_0) \\ &\leq C\left(\|\boldsymbol{u}_0\|_{0,2}^2 + \|\boldsymbol{u}_0\|_{0,3}^3 + \|\boldsymbol{g}\|_{0,2}^2\right) \\ &\leq C\left(\|f\|_{0,2}^2 + \|f\|_{0,2}^3 + \|\boldsymbol{g}\|_{0,2}^2\right). \end{aligned}$$

The result (19) follows immediately by setting $\varepsilon = \frac{1}{2}$.

Using once again the inf-sup condition Lemma 3.2 and Lemma 3.1, we obtain the existence of p, and $(\boldsymbol{u}, p)$ satisfies

$$\|p\|_{0,2} \leq \frac{1}{\gamma}\frac{(p, \nabla \cdot \boldsymbol{v})}{\|\boldsymbol{v}\|_X} = \frac{1}{\gamma}\frac{(A(\boldsymbol{u}), \boldsymbol{v}) - (\boldsymbol{g}, \boldsymbol{v})}{\|\boldsymbol{v}\|_X} \leq C\left(\|\boldsymbol{u}\|_{0,2} + \|\boldsymbol{u}\|_{0,3}^2 + \|\boldsymbol{g}\|_{0,2}\right).$$

Using some inequality techniques, the result (20) follows immediately. □

Lemma 3.4 (Inf-sup condition of the discrete problem) *There exists a positive constant γ_h independent of h such that*

$$\inf_{w_h\in M_h}\sup_{\boldsymbol{v}_h\in X_h}\frac{(w_h,\nabla\cdot\boldsymbol{v}_h)}{\|w_h\|_{M_h}\|\boldsymbol{v}_h\|_{X_h}}\geq\tilde{\gamma}. \tag{24}$$

Proof Let $w_h\in M_h$ and $\boldsymbol{v}^*=-\frac{2(\nabla p^*-\boldsymbol{g})}{\frac{\mu}{\rho K}+\sqrt{(\frac{\mu}{\rho K})^2+4\frac{\beta}{\rho}|\nabla p^*-\boldsymbol{g}|}}$, where p^* is the solution of the following problem:

$$\begin{cases}-\nabla\cdot\left(\frac{2(\nabla p^*-\boldsymbol{g})}{\frac{\mu}{\rho K}+\sqrt{(\frac{\mu}{\rho K})^2+4\frac{\beta}{\rho}|\nabla p^*-\boldsymbol{g}|}}\right)=w_h, & \text{in }\Omega,\\ p^*=0, & \text{on }\partial\Omega.\end{cases}$$

This problem has a unique solution $p^*\in W_0^{1,\frac{3}{2}}(\Omega)$ and there exists a constant $C>0$, such that $\|p^*-Z\|_{1,\frac{3}{2}}\leq C\|w_h\|_{0,2}^2$. Thus $\boldsymbol{v}^*\in X$ and $\|\boldsymbol{v}^*\|_X\leq C\|w_h\|_{0,2}$. Now, let $\boldsymbol{v}=\nabla p$, p is the solution of the Dirichlet problem

$$\begin{cases}-\triangle p=\nabla\cdot\boldsymbol{v}^*, & \text{in }\Omega,\\ p=0, & \text{on }\partial\Omega.\end{cases}$$

Since $\nabla\cdot\boldsymbol{v}\in L^2(\Omega)$, we have $\boldsymbol{v}\in(H^1(\Omega))^d$, there exists a constant C such that

$$\|\boldsymbol{v}\|_{0,3}\leq C\|\boldsymbol{v}\|_{1,2}\leq C\|\nabla\cdot\boldsymbol{v}^*\|_{0,2},$$

and $\nabla\cdot\boldsymbol{v}=\nabla\cdot\boldsymbol{v}^*$ in Ω.

Finally, let $\boldsymbol{v}_h=\Pi_h\boldsymbol{v}$, and we have

$$\int_\Omega\nabla\cdot\boldsymbol{v}_h w_h\,\mathrm{d}\boldsymbol{x}=\int_\Omega\nabla\cdot\boldsymbol{v}w_h\,\mathrm{d}\boldsymbol{x}=\int_\Omega\nabla\cdot\boldsymbol{v}^*w_h\,\mathrm{d}\boldsymbol{x}=\|w_h\|_{0,2}^2$$

and

$$\|\boldsymbol{v}_h\|_X\leq C(\|\boldsymbol{v}\|_{0,3}+\|\nabla\cdot\boldsymbol{v}\|_{0,2})\leq C\|\nabla\cdot\boldsymbol{v}^*\|_{0,2}\leq C\|w_h\|_{0,2}.$$

Therefore, discrete inf-sup condition (24) follows immediately. □

Theorem 3.5 (Existence and uniqueness of the discrete problem) *The problem (P_h) has a unique solution $(\boldsymbol{v}_h,w_h)\in X_h\times M_h$, and there exists a constant $C>0$ such that*

$$\|\boldsymbol{u}_h\|_{0,2}^2+\|\boldsymbol{u}_h\|_{0,3}^3+\|\nabla\cdot\boldsymbol{u}_h\|_{0,2}^2\leq C\big(\|f\|_{0,2}^2+\|f\|_{0,2}^3+\|\boldsymbol{g}\|_{0,2}^2\big), \tag{25}$$

$$\|p_h\|_{0,2}\leq C\big(\|f\|_{0,2}+\|f\|_{0,2}^2+\|\boldsymbol{g}\|_{0,2}+\|\boldsymbol{g}\|_{0,2}^2\big). \tag{26}$$

Proof Using the discrete inf-sup condition, i.e. Lemma 3.4, the existence and uniqueness is easy to verified, which is similar to Theorem 3.3. □

4 Error Estimate

In this section we will give the error evaluation between the analytical solution and the approximation solution.

Subtracting the weak equation and its finite element approximation, error equation is deserved as follows:

$$(A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{v}_h) - (p - p_h, \nabla \cdot \boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in X_h,$$

$$(w_h, \nabla \cdot (\boldsymbol{u} - \boldsymbol{u}_h)) = 0, \quad \forall w_h \in M_h.$$

Using the interpolation or projection (6), (9) and (12), we get

$$(w_h, \nabla \cdot (\Pi_h \boldsymbol{u} - \boldsymbol{u}_h)) = -(w_h, \nabla \cdot (\boldsymbol{u} - \Pi_h \boldsymbol{u})) = 0, \quad \forall w_h \in M_h,$$

$$(A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{v}_h) - (P_h p - p_h, \nabla \cdot \boldsymbol{v}_h) = (p - P_h p, \nabla \cdot \boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in X_h.$$

Set $\boldsymbol{v}_h = \Pi_h \boldsymbol{u} - \boldsymbol{u}_h$, and using the fact that

$$(P_h p - p_h, \nabla \cdot (\Pi_h \boldsymbol{u} - \boldsymbol{u}_h)) = 0,$$

we get

$$(A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \Pi_h \boldsymbol{u} - \boldsymbol{u}_h) = 0,$$

i.e.

$$(A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{u} - \boldsymbol{u}_h) = (A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{u} - \Pi_h \boldsymbol{u}). \tag{27}$$

We will use the interpolant error $\boldsymbol{u} - \Pi_h \boldsymbol{u}$ to control $\boldsymbol{u} - \boldsymbol{u}_h$. To control $P_h p - p_h$, we will use the following equality

$$(P_h p - p_h, \nabla \cdot \boldsymbol{v}_h) = (A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{v}_h) \tag{28}$$

and the Inf-sup condition (24).

The next two lemmas of the absolute value type monotone operator are used in the error estimate, they can be found in [14, 22].

Lemma 4.1 *Let $\boldsymbol{x}, \boldsymbol{h} \in R^d$. $\boldsymbol{f} : R^d \to R^d$ is defined as $\boldsymbol{f}(\boldsymbol{x}) = |\boldsymbol{x}|\boldsymbol{x}$. We have*

$$C_h^1(|\boldsymbol{x}| + |\boldsymbol{x} + \boldsymbol{h}|)|\boldsymbol{h}|^2 \le (\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{h}), \tag{29}$$

$$|\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x})| \le C_h^2(|\boldsymbol{x}| + |\boldsymbol{x} + \boldsymbol{h}|)|\boldsymbol{h}|, \tag{30}$$

$$C_h^3|\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x})||\boldsymbol{h}| \le (\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{h}). \tag{31}$$

Proof First, we differentiate $\boldsymbol{f}(\boldsymbol{x})$,

$$\frac{\partial f_i}{\partial x_j}(\boldsymbol{x}) = (x_i x_j + |\boldsymbol{x}|^2 \delta_{ij})|\boldsymbol{x}|^{-1}.$$

Then, for $\boldsymbol{k} = (k_1, k2, \ldots, k_d) \in R^d$, by Taylor expansion we get,

$$(\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{k}) = \sum_{i=1}^{d} (f_i(\boldsymbol{x}+\boldsymbol{h}) - f_i(\boldsymbol{x}))k_i = \sum_{i=1}^{d}\sum_{j=1}^{d} \int_0^1 \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}+t\boldsymbol{h})h_j k_i \,\mathrm{d}t.$$

Let $\boldsymbol{x}^t = \boldsymbol{x} + t\boldsymbol{h}$,

$$\begin{aligned}(\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{k}) &= \sum_{i=1}^{d}\sum_{j=1}^{d} \int_0^1 (x_i^t x_j^t + |\boldsymbol{x}^t|^2 \delta_{ij})|\boldsymbol{x}^t|^{-1} h_j k_i \,\mathrm{d}t \\ &= \sum_{i=1}^{d} \int_0^1 |\boldsymbol{x}^t| h_i k_i \,\mathrm{d}t + \sum_{i=1}^{d}\sum_{j=1}^{d} \int_0^1 x_i^t x_j^t |x^t|^{-1} h_j k_i \,\mathrm{d}t \\ &= (\boldsymbol{h}, \boldsymbol{k}) \int_0^1 |\boldsymbol{x}^t| \,\mathrm{d}t + \int_0^1 |\boldsymbol{x}^t|^{-1} (\boldsymbol{x}^t, \boldsymbol{h})(\boldsymbol{x}^t, \boldsymbol{k}) \,\mathrm{d}t.\end{aligned}$$

For the seek of proving (29), we let $\boldsymbol{k} = \boldsymbol{h}$, and get

$$(\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{h}) \geq |\boldsymbol{h}|^2 \int_0^1 |\boldsymbol{x}^t| \,\mathrm{d}t.$$

Next we estimate $\int_0^1 |\boldsymbol{x}^t| \,\mathrm{d}t$. It is clear that, $\forall \beta > 0$, there exists $C_{1,d,\beta} > 0$, $C_{2,d,\beta} > 0$, such that

$$C_{1,d,\beta} \sum_{i=1}^{d} |x_i| \leq \left(\sum_{i=1}^{d} |x_i|^\beta \right)^{\frac{1}{\beta}} \leq C_{2,d,\beta} \sum_{i=1}^{d} |x_i|.$$

Thus

$$\begin{aligned}\int_0^1 |\boldsymbol{x}^t| \,\mathrm{d}t &= \int_0^1 \left(\sum_{i=1}^{d} (x_i + t h_i)^2 \right)^{\frac{1}{2}} \mathrm{d}t \geq C_{1,d,2} \int_0^1 \sum_{i=1}^{d} |x_i + t h_i| \,\mathrm{d}t \\ &\geq C_{1,d,2} \left(\frac{1}{2} \sum_{i=1}^{d} |x_i| \int_0^1 \left| 1 + t \frac{h_i}{x_i} \right| \mathrm{d}t \right. \\ &\quad \left. + \frac{1}{2} \sum_{i=1}^{d} |x_i + h_i| \int_0^1 \left| 1 + (t-1) \frac{h_i}{x_i + h_i} \right| \mathrm{d}t \right) \\ &\geq C_{1,d,2} C_{2,d,2}^{-1} \left[\frac{1}{2} |\boldsymbol{x}| \inf_{\lambda \in R} \int_0^1 |1 + t\lambda| \,\mathrm{d}t + \frac{1}{2} |\boldsymbol{x} + \boldsymbol{h}| \inf_{\lambda \in R} \int_0^1 |1 + (t-1)\lambda| \,\mathrm{d}t \right].\end{aligned}$$

Since $\phi(\lambda)=\int_0^1|1+t\lambda|\,\mathrm{d}t$ and $\psi(\lambda)=\int_0^1|1+(t-1)\lambda|\,\mathrm{d}t$ are continuous positive definite, and $\phi(\lambda)$ and $\psi(\lambda)\to+\infty$ as $|\lambda|\to+\infty$. It's easy to see that they have minimum value $\frac{1}{2}$ at $\lambda=-2$, and $\lambda=2$ respectively. Thus, there exist a constant C_d such that

$$\int_0^1|\boldsymbol{x}^t|\,\mathrm{d}t\geq C_{2,d,\frac{1}{2}}^{-1}C_{1,d,\frac{1}{2}}C_d(|\boldsymbol{x}|+|\boldsymbol{x}+\boldsymbol{h}|)\geq C_{0,d}^1(|\boldsymbol{x}|+|\boldsymbol{x}+\boldsymbol{h}|).$$

Therefore inequality (29) holds.

Now we prove (30)

$$|(\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{x}),\boldsymbol{k})|\leq|\boldsymbol{h}||\boldsymbol{k}|2\int_0^1|\boldsymbol{x}^t|\,\mathrm{d}t.$$

Therefore

$$|\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{x})|\leq2|\boldsymbol{h}|\int_0^1|\boldsymbol{x}^t|\,\mathrm{d}t.$$

Since $|\boldsymbol{x}+t\boldsymbol{h}|\leq|\boldsymbol{x}|+|\boldsymbol{h}|=|\boldsymbol{x}|+|-\boldsymbol{x}+\boldsymbol{x}+\boldsymbol{h}|\leq2|\boldsymbol{x}|+|\boldsymbol{x}+\boldsymbol{h}|$. We get inequality (30).

It is easy to get inequality (31) from (29) and (30). □

Lemma 4.2 *Let $\boldsymbol{f}(\boldsymbol{v})=|\boldsymbol{v}|\boldsymbol{v}$. There exist positive constants C_i, $i=1,2,3,4$, such that for $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in L^3(\Omega)^d$, there hold*

$$C_1\int_\Omega(|\boldsymbol{u}|+|\boldsymbol{v}|)|\boldsymbol{v}-\boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{x}\leq\int_\Omega(\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}\boldsymbol{x},\tag{32}$$

$$\int_\Omega(\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{w})\,\mathrm{d}\boldsymbol{x}\leq C_2\left[\int_\Omega(|\boldsymbol{u}|+|\boldsymbol{v}|)|\boldsymbol{v}-\boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{x}\right]^{\frac{1}{2}}\left[\|\boldsymbol{u}\|_{0,3}^{\frac{1}{2}}+\|\boldsymbol{v}\|_{0,3}^{\frac{1}{2}}\right]\|\boldsymbol{w}\|_{0,3},\tag{33}$$

$$C_3\|\boldsymbol{v}-\boldsymbol{u}\|_{0,3}^3\leq\int_\Omega(\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}\boldsymbol{x},\tag{34}$$

$$C_4\int_\Omega|\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u})||\boldsymbol{v}-\boldsymbol{u}|\,\mathrm{d}\boldsymbol{x}\leq\int_\Omega(\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}\boldsymbol{x}.\tag{35}$$

Proof Using (29), we easily get

$$\int_\Omega(\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}\boldsymbol{x}\geq C\int_\Omega(|\boldsymbol{v}|+|\boldsymbol{u}|)|\boldsymbol{v}-\boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{x},$$

which completes the proof of (32).

Using the inequality (32) and the Minkowski inequality, we easily get

$$\int_\Omega(\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}\boldsymbol{x}\geq C\|\boldsymbol{v}-\boldsymbol{u}\|_{0,3}^3.$$

which completes the proof of (34).

Now we prove (33). Using the inequality (30),

$$\begin{aligned}\int_\Omega (\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{w})\,\mathrm{d}\boldsymbol{x} &\le \int_\Omega |\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u})||\boldsymbol{w}|\,\mathrm{d}\boldsymbol{x}\\ &\le C\int_\Omega (|\boldsymbol{v}|+|\boldsymbol{u}|)|\boldsymbol{v}-\boldsymbol{u}||\boldsymbol{w}|\,\mathrm{d}\boldsymbol{x}\\ &= C\int_\Omega (|\boldsymbol{v}|+|\boldsymbol{u}|)^{\frac{1}{2}}|\boldsymbol{v}-\boldsymbol{u}|(|\boldsymbol{v}|+|\boldsymbol{u}|)^{\frac{1}{2}}|\boldsymbol{w}|\,\mathrm{d}\boldsymbol{x}\\ &\le C\left[\int_\Omega (|\boldsymbol{v}|+|\boldsymbol{u}|)|\boldsymbol{v}-\boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{x}\right]^{\frac{1}{2}}\left[\int_\Omega (|\boldsymbol{v}|+|\boldsymbol{u}|)|\boldsymbol{w}|^2\,\mathrm{d}\boldsymbol{x}\right]^{\frac{1}{2}}.\end{aligned}$$

To control the second term of the right hand side, let $\gamma = 3\gamma' = \frac{1}{1-\frac{1}{\gamma}} = \frac{3}{2}$, we get

$$\begin{aligned}\int_\Omega (|\boldsymbol{v}|+|\boldsymbol{u}|)|\boldsymbol{w}|^2\,\mathrm{d}\boldsymbol{x} &\le \left[\int_\Omega (|\boldsymbol{v}|+|\boldsymbol{u}|)^{\gamma}\,\mathrm{d}\boldsymbol{x}\right]^{\frac{1}{\gamma}}\left[\int_\Omega |\boldsymbol{w}|^{2\gamma'}\,\mathrm{d}\boldsymbol{x}\right]^{\frac{1}{\gamma'}}\\ &\le C\left[\int_\Omega (|\boldsymbol{v}|^3+|\boldsymbol{u}|^3)\,\mathrm{d}\boldsymbol{x}\right]^{\frac{1}{3}}\left[\int_\Omega |\boldsymbol{w}|^3\,\mathrm{d}\boldsymbol{x}\right]^{\frac{2}{3}}\\ &\le C(\|\boldsymbol{v}\|_{0,3}+\|\boldsymbol{u}\|_{0,3})\|\boldsymbol{w}\|_{0,3}^2.\end{aligned}$$

Using the (31), we can get (35) easily,

$$\int_\Omega (\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}\boldsymbol{x} \ge C\int_\Omega |\boldsymbol{f}(\boldsymbol{v})-\boldsymbol{f}(\boldsymbol{u})||\boldsymbol{v}-\boldsymbol{u}|\,\mathrm{d}\boldsymbol{x}.$$ □

Theorem 4.3 *Let* $(\boldsymbol{u},p)\in X\times M$ *be the solution of the weak problem* (4) *or* (15), *and* $(\boldsymbol{u}_h,p_h)\in X_h\times M_h$ *be the solution of the discrete problem* (5) *or* (16), *then there exists a constant* C *independent of* h *such that*

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,2}^2+\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,3}^3 \le C\left\{\|\boldsymbol{u}-\Pi_h\boldsymbol{u}\|_{0,2}^2+\|\boldsymbol{u}-\Pi_h\boldsymbol{u}\|_{0,3}^2\right\}, \tag{36}$$

$$\|p-p_h\|_{0,2} \le C\left\{\|\boldsymbol{u}-\Pi_h\boldsymbol{u}\|_{0,2}+\|\boldsymbol{u}-\Pi_h\boldsymbol{u}\|_{0,3}+\|p-P_hp\|_{0,2}\right\}. \tag{37}$$

Proof We consider the left hand side and right hand side of the error equation (27),

$$\begin{aligned}l.h.s. &= \int_\Omega (A(\boldsymbol{u})-A(\boldsymbol{u}_h))\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\,\mathrm{d}\boldsymbol{x}\\ &= \int_\Omega \frac{\mu}{\rho}K^{-1}(\boldsymbol{u}-\boldsymbol{u}_h)\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\,\mathrm{d}\boldsymbol{x}+\left(\frac{1}{2}+\frac{1}{2}\right)\int_\Omega (|\boldsymbol{u}|\boldsymbol{u}-|\boldsymbol{u}_h|\boldsymbol{u}_h)\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\,\mathrm{d}\boldsymbol{x}\\ &\ge C_0\left\{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,2}^2+\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,3}^3+\int_\Omega (|\boldsymbol{u}|+|\boldsymbol{u}_h|)|\boldsymbol{u}-\boldsymbol{u}_h|^2\,\mathrm{d}\boldsymbol{x}\right\},\end{aligned}$$

$$
\begin{aligned}
r.h.s. &= \int_\Omega (A(\boldsymbol{u}) - A(\boldsymbol{u}_h)) \cdot (\boldsymbol{u} - \Pi_h \boldsymbol{u})\,\mathrm{d}\boldsymbol{x} \\
&= \int_\Omega \frac{\mu}{\rho} K^{-1}(\boldsymbol{u} - \boldsymbol{u}_h) \cdot (\boldsymbol{u} - \Pi_h \boldsymbol{u})\,\mathrm{d}\boldsymbol{x} + \int_\Omega (|\boldsymbol{u}|\boldsymbol{u} - |\boldsymbol{u}_h|\boldsymbol{u}_h) \cdot (\boldsymbol{u} - \Pi_h \boldsymbol{u})\,\mathrm{d}\boldsymbol{x} \\
&\le C \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,2} \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,2} \\
&\quad + C \left[\int_\Omega (|\boldsymbol{u}| + |\boldsymbol{u}_h|) |\boldsymbol{u} - \boldsymbol{u}_h|^2\,\mathrm{d}\boldsymbol{x} \right]^{\frac{1}{2}} \times \left[\|\boldsymbol{u}\|_{0,3}^{\frac{1}{2}} + \|\boldsymbol{u}_h\|_{0,3}^{\frac{1}{2}} \right] \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,3} \\
&\le \varepsilon \left\{ \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,2}^2 + \int_\Omega (|\boldsymbol{u}| + |\boldsymbol{u}_h|) |\boldsymbol{u} - \boldsymbol{u}_h|^2\,\mathrm{d}\boldsymbol{x} \right\} \\
&\quad + \frac{C_5}{\varepsilon} \left\{ \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,2}^2 + \left[\|\boldsymbol{u}\|_{0,3}^{\frac{1}{2}} + \|\boldsymbol{u}_h\|_{0,3}^{\frac{1}{2}} \right]^2 \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,3}^2 \right\}.
\end{aligned}
$$

Here C_0 and C_5 are positive constants independent of h and ε. Choosing $\varepsilon = C_0/2$ and applying the above inequalities and (27), we deserve

$$
\begin{aligned}
&\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,2}^2 + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,3}^3 + \int_\Omega (|\boldsymbol{u}| + |\boldsymbol{u}_h|) |\boldsymbol{u} - \boldsymbol{u}_h|^2\,\mathrm{d}\boldsymbol{x} \\
&\quad \le C \{ \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,2}^2 + \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,3}^2 \}.
\end{aligned}
$$

Dropping the positive term, we get the inequality (36).

Using the following inequalities

$$
\begin{aligned}
&\int_\Omega (A(\boldsymbol{u}) - A(\boldsymbol{u}_h)) \cdot \boldsymbol{v}_h\,\mathrm{d}\boldsymbol{x} \\
&\quad \le C \left\{ \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,2} \|\boldsymbol{v}\|_{0,2} + \left[\int_\Omega (|\boldsymbol{u}| + |\boldsymbol{u}_h|) |\boldsymbol{u} - \boldsymbol{u}_h|^2\,\mathrm{d}\boldsymbol{x} \right]^{\frac{1}{2}} \right. \\
&\qquad \left. \times \left[\|\boldsymbol{u}\|_{0,3}^{\frac{1}{2}} + \|\boldsymbol{u}_h\|_{0,3}^{\frac{1}{2}} \right] \|\boldsymbol{v}_h\|_{0,3} \right\},
\end{aligned}
$$

$$
\begin{aligned}
\gamma_h \|P_h p - p_h\|_{0,2} &\le \sup_{\boldsymbol{v}_h \in X_h} \frac{(P_h p - p_h, \nabla \cdot \boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_X} = \sup_{\boldsymbol{v}_h \in X_h} \frac{(A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_X} \\
&\le C \left\{ \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,2} + \left[\int_\Omega (|\boldsymbol{u}| + |\boldsymbol{u}_h|) |\boldsymbol{u} - \boldsymbol{u}_h|^2\,\mathrm{d}\boldsymbol{x} \right]^{\frac{1}{2}} \right. \\
&\qquad \left. \times \left[\|\boldsymbol{u}\|_{0,3}^{\frac{1}{2}} + \|\boldsymbol{u}_h\|_{0,3}^{\frac{1}{2}} \right] \right\} \\
&\le C \left\{ \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,2} + \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_{0,3} \right\},
\end{aligned}
$$

and

$$
\|p - p_h\|_{0,2} \le \|p - P_h p\|_{0,2} + \|P_h p - p_h\|_{0,2},
$$

it is easy to get the inequality (37). □

Theorem 4.4 *Let* $(\boldsymbol{u}, p) \in X \times M$ *be the solution of weak problem* (4) *or* (15), *and* $(\boldsymbol{u}_h, p_h) \in X_h \times M_h$ *be the solution of discrete problem* (5) *or* (16). *If* $(\boldsymbol{u}, p) \in W^{s,3}(\Omega)^d \times W^{s,\frac{3}{2}}(\Omega)$, *then there exists a constant* C *independent of* h *such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,2}^2 + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,3}^3 \leq Ch^{2s}, \quad 1 \leq s \leq k+1, \tag{38}$$

$$\|p - p_h\|_{0,2} \leq Ch^s, \quad 1 \leq s \leq k+1. \tag{39}$$

Proof It is easy to get the estimates (38) and (39) by substituting the interpolation error inequalities (7) and (10) into the inequalities (36) and (37) of Theorem 4.3. □

5 An Iterative Algorithm

In this section, an iterative algorithm is proposed for practical implementation. We start by introducing some symbols to illustrate the matrix form of the mixed element approximation. Let $\tilde{u}$, $\tilde{p}$ be the coefficients of the global basis of the finite element approximation, whose dimension is m_1, m_2 respectively, i.e.

$$\boldsymbol{u}_h = \sum_{i=1}^{m_1} \tilde{u}^i \boldsymbol{\phi}_i, \qquad p_h = \sum_{i=1}^{m_2} \tilde{p}^i \psi_i.$$

Where $\tilde{u} = [\tilde{u}^1, \tilde{u}^2, \ldots, \tilde{u}^{m_1}]^T$, $\tilde{p} = [\tilde{p}^1, \tilde{p}^2, \ldots, \tilde{p}^{m_2}]^T$ and $\boldsymbol{\phi}$ is the basis of velocity, ψ is the basis of pressure.

Let $D^{-1}(\tilde{u})$, B be the functional matrices. $D^{-1}(\tilde{u})\tilde{u}$ is a nonlinear vector function of $\tilde{u}$ corresponding to

$$\int_\Omega \left(\frac{\mu}{\rho} K^{-1} \boldsymbol{u}_h + \frac{\beta}{\rho} |\boldsymbol{u}_h| \boldsymbol{u}_h \right) \cdot \boldsymbol{v}_h \, \mathrm{d}\boldsymbol{x}.$$

B is the functional matrix and B, B^T are corresponding to

$$-\int_\Omega p_h \nabla \cdot \boldsymbol{v}_h \, \mathrm{d}\boldsymbol{x}, \qquad -\int_\Omega w_h \nabla \cdot \boldsymbol{u}_h \, \mathrm{d}\boldsymbol{x},$$

respectively. G and F are functional vectors corresponding to

$$-\int_{\partial\Omega} f_D \boldsymbol{v}_h \cdot \boldsymbol{n} \, \mathrm{d}s + \int_\Omega \boldsymbol{g} \cdot \boldsymbol{v}_h \, \mathrm{d}\boldsymbol{x}, \qquad -\int_\Omega w_h f \, \mathrm{d}\boldsymbol{x},$$

respectively.

We first list the matrix form of mixed element approximation (5).

$$\begin{cases} \text{Find } \tilde{u} \in R^{m_1}, \tilde{p} \in R^{m_2} \text{ such that,} \\ \begin{pmatrix} D^{-1}(\tilde{u}) & B \\ B^T & O \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{p} \end{pmatrix} = \begin{pmatrix} G \\ F \end{pmatrix}. \end{cases} \tag{40}$$

The system is symmetric under the condition that K is scalar or a symmetric tensor. It is easy to see the system is also nonlinear and indefinite, which are the reasons for designing iterative algorithm for solving the problem (40). We give the iterative algorithm as below.

Iterative Algorithm of Mixed Element Approximation

$$\begin{cases} \text{Given arbitrary } \boldsymbol{u}_h^0 \in X, \text{ find } \boldsymbol{u}_h^{n+1} \in X,\ p_h^{n+1} \in M \text{ such that,} \\ (D^{-1}(\boldsymbol{u}_h^n)\boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) - (p_h^{n+1}, \nabla \cdot \boldsymbol{v}_h) = (\boldsymbol{g}, \boldsymbol{v}_h) + (f_D, \boldsymbol{v}_h \cdot \boldsymbol{n}), \quad \forall \boldsymbol{v}_h \in X_h, \\ -(w_h, \nabla \cdot \boldsymbol{u}_h^{n+1}) = -(w_h, f), \quad \forall w_h \in M_h. \end{cases} \tag{41}$$

Matrix form of the Iterative Algorithm of Mixed Element Approximation

$$\begin{cases} \text{Given } \tilde{u}^0 \in R^{m_1}, \text{ find } \tilde{u}^{n+1} \in R^{m_1},\ \tilde{p}^{n+1} \in R^{m_2} \text{ such that,} \\ \begin{pmatrix} D^{-1}(\tilde{u}^n) & B \\ B^T & O \end{pmatrix} \begin{pmatrix} \tilde{u}^{n+1} \\ \tilde{p}^{n+1} \end{pmatrix} = \begin{pmatrix} G \\ F \end{pmatrix}. \end{cases} \tag{42}$$

This algorithm is linear, thus algebraic iterative algorithm like Newton type is not necessary, which reduces the implementation difficulty.

Although we do some efforts,the convergence analysis is still under consideration.

6 Numerical Experiment

6.1 Numerical Examples

In this section we carry out numerical experiments using the lowest order Raviart-Thomas finite element; we test several examples using triangular and rectangular element in two-dimension domain to verify the rates of convergence. For simplicity, the region of examples are unit square, i.e. $\Omega = [0, 1] \times [0, 1]$. The boundary condition is $f_D = 0$ on $\partial\Omega$. The permeability is a constant, and the viscosity and density are also constants. For simplicity, take $\mu = \rho = K = 1$.

The polygonalization of rectangle in 2D is uniform subdivision in each dimension. The subdivision scale is characterized by *scale*. For the quasi-uniform polygonalization, *scale* is the same order of magnitude of the discretization parameter h. For example *scale* $= [0.1, 0.1]$ in 2D is corresponding to the case of $h = 1/10$. The triangulation in 2D can be based on the polygonalization of rectangle in 2D or generates using Delaunay triangulation. We refer to [2] for the implementation details of lowest order Raviart-Thomas mixed element.

The numerical examples constructed in this section can be divided into two categories. First, Examples 1 to 4, in which we know the analytic solution of the problem, are used to justify the convergence rate of the method proposed in this paper. In Sect. 6.2 Examples 1 to 4 are carried out to show the convergence of the iterate algorithm, while in Sect. 6.3 Examples 1 to 4 are experimented to demonstrate the convergence rate of the mixed element approximation. Examples 1 to 4 are listed as below.

Example 1 (Forchheimer problem without source)

$$\begin{cases} p(x,y) = (x-x^2)(y-y^2), \qquad \boldsymbol{u}(x,y) = (\exp(x)\sin y, \exp(x)\cos y)^T, \\ f(x,y) = 0, \\ g(x,y) = \left(\dfrac{\mu}{\rho K} + \dfrac{\beta}{\rho}\sqrt{(\exp(x)\sin y)^2 + (\exp(x)\cos y)^2}\right)(\exp(x)\sin y, \exp(x)\cos y)^T \\ \qquad\quad + ((1-2x)(y-y^2), (x-x^2)(1-2y))^T. \end{cases}$$

Example 2 (Forchheimer problem without source)

$$\begin{cases} p(x,y) = \sin\xi_1 x \sin\xi_2 y, \qquad \boldsymbol{u}(x,y) = (\exp(x)\sin y, \exp(x)\cos y)^T, \\ f(x,y) = 0, \\ g(x,y) = \left(\dfrac{\mu}{\rho K} + \dfrac{\beta}{\rho}\sqrt{(\exp(x)\sin y)^2 + (\exp(x)\cos y)^2}\right)(\exp(x)\sin y, \exp(x)\cos y)^T \\ \qquad\quad + (\xi_1\cos\xi_1 x \sin\xi_2 y, \xi_2 \sin\xi_1 x \cos\xi_2 y)^T, \quad \xi_1 = \pi, \xi_2 = \pi. \end{cases}$$

Example 3 (Forchheimer problem with source)

$$\begin{cases} p(x,y) = (x-x^2)(y-y^2), \qquad \boldsymbol{u}(x,y) = (x\exp(y), y\exp(x))^T, \\ f(x,y) = \exp(x) + \exp(y), \\ g(x,y) = \left(\dfrac{\mu}{\rho K} + \dfrac{\beta}{\rho}\sqrt{(x\exp(y))^2 + (y\exp(x))^2}\right)(x\exp(y), y\exp(x))^T \\ \qquad\quad + ((1-2x)(y-y^2), (x-x^2)(1-2y))^T. \end{cases}$$

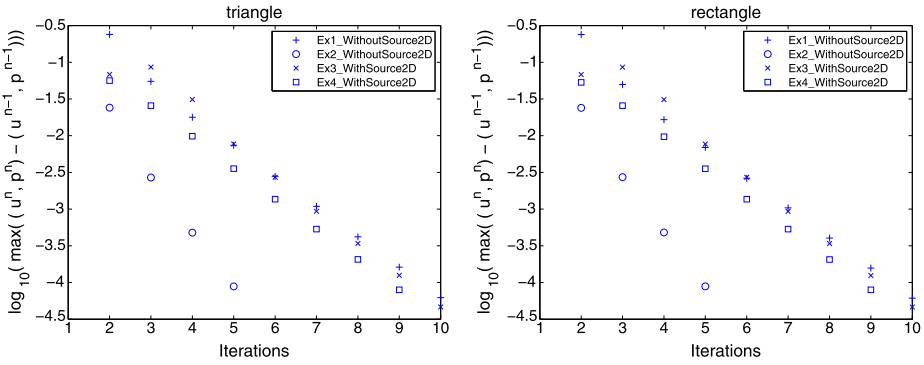
Example 4 (Forchheimer problem with source)

$$\begin{cases} p(x,y) = \sin\xi_1 x \sin\xi_2 y, \qquad \boldsymbol{u}(x,y) = (x\exp(y), y\exp(x))^T, \\ f(x,y) = \exp(x) + \exp(y), \\ g(x,y) = \left(\dfrac{\mu}{\rho K} + \dfrac{\beta}{\rho}\sqrt{(x\exp(y))^2 + (y\exp(x))^2}\right)(x\exp(y), y\exp(x))^T \\ \qquad\quad + (\xi_1\cos\xi_1 x \sin\xi_2 y, \xi_2 \sin\xi_1 x \cos\xi_2 y)^T, \quad \xi_1 = \pi, \xi_2 = \pi. \end{cases}$$

Second, Examples 5 to 8 are used to observe the different behavior of the Forchheimer flow and Darcy flow. These examples are constructed by using the same right hand side and boundary condition for the different operators: the Darcy and the Forchheimer operator, so the different behaviors are just caused by the operator. We show the different behaviors of the Forchheimer flow and Darcy flow in Sect. 6.4. The needed source terms and boundary condition are listed as below.

Example 5 (Forchheimer and Darcy problems without source)

$$\begin{cases} f(x,y) = 0, \\ g(x,y) = (\sin\pi x, \sin\pi y)^T. \end{cases}$$



(a) Triangular triangulation (b) Rectangular triangulation

Fig. 1 Convergence of iterative algorithm

Example 6 (Forchheimer and Darcy problems without source)

$$\begin{cases} f(x,y) = 0, \\ g(x,y) = (x - x^2, y - y^2)^T. \end{cases}$$

Example 7 (Forchheimer and Darcy problems with source)

$$\begin{cases} f(x,y) = \exp\{-[(x - \frac{1}{2})^2 + (y - \frac{1}{2})^2]\}, \\ g(x,y) = (1,0)^T. \end{cases}$$

Example 8 (Forchheimer and Darcy problems with source)

$$\begin{cases} f(x,y) = \sin\xi_1 x \sin\xi_2 y, \quad \xi_1 = \pi,\ \xi_2 = 2\pi, \\ g(x,y) = (1,0)^T. \end{cases}$$

6.2 Convergence of Iterative Algorithm

In this section, we will show the convergence of the iterative algorithm by some numerical examples. We use the maximum norm in error control of the iterative algorithm, i.e.

$$eps = \max\left(\left|\begin{pmatrix} \tilde{u}^{n+1} \\ \tilde{p}^{n+1} \end{pmatrix} - \begin{pmatrix} \tilde{u}^{n} \\ \tilde{p}^{n} \end{pmatrix}\right|\right).$$

We test Examples 1 to 4 to obtain the convergence of iterative algorithm. The parameters are set as bellow. The discretization parameter $h = 1/20$. The iterative algorithm error control is $eps = 10^{-4}$. The Forchheimer number is set to be 1, i.e. $\beta = 1$. The results are listed in Fig. 1.

Table 2 Relative error for triangular triangulation

	Resolution	$E_{0,2}^{p,h}$	$E_{0,\infty}^{p,h}$	$E_{0,2}^{\boldsymbol{u},h}$	$E_{0,3}^{\boldsymbol{u},h}$	$E_{0,\infty}^{\boldsymbol{u},h}$	$E_{\mathrm{div},2}^{\boldsymbol{u},h}$
Ex1	$h=1/5$	4.54E−1	9.61E−1	9.00E−2	9.16E−2	1.71E−1	–
	$h=1/10$	1.58E−1	3.94E−1	4.57E−2	4.74E−2	9.27E−2	–
	$h=1/20$	6.13E−2	1.61E−1	2.29E−2	2.39E−2	4.77E−2	–
	$h=1/40$	2.75E−2	7.12E−2	1.15E−2	1.20E−2	2.41E−2	–
	$h=1/80$	1.33E−2	3.34E−2	5.76E−3	6.02E−3	1.22E−2	–
	$h=1/160$	6.61E−3	1.61E−2	2.90E−3	3.03E−3	6.41E−3	–
Ex2	$h=1/5$	3.23E−1	7.18E−1	9.00E−2	9.16E−2	1.71E−1	–
	$h=1/10$	1.64E−1	3.80E−1	4.57E−2	4.74E−2	9.27E−2	–
	$h=1/20$	8.26E−2	1.95E−1	2.29E−2	2.39E−2	4.77E−2	–
	$h=1/40$	4.13E−2	9.82E−2	1.15E−2	1.20E−2	2.41E−2	–
	$h=1/80$	2.06E−2	4.91E−2	5.76E−3	6.02E−3	1.22E−2	–
	$h=1/160$	1.03E−2	2.45E−2	2.90E−3	3.03E−3	6.41E−3	–
Ex3	$h=1/5$	3.02E−1	6.31E−1	6.03E−2	4.54E−2	8.92E−2	4.10E−2
	$h=1/10$	1.21E−1	2.89E−1	3.03E−2	2.36E−2	4.58E−2	2.05E−2
	$h=1/20$	5.48E−2	1.36E−1	1.52E−2	1.19E−2	2.29E−2	1.02E−2
	$h=1/40$	2.66E−2	6.54E−2	7.61E−3	6.02E−3	1.22E−2	5.13E−3
	$h=1/80$	1.32E−2	3.19E−2	3.80E−3	3.05E−3	6.68E−3	2.56E−3
	$h=1/160$	6.59E−3	1.57E−2	1.90E−3	1.56E−3	4.02E−3	1.28E−3
Ex4	$h=1/5$	3.22E−1	7.19E−1	6.03E−2	4.54E−2	8.92E−2	4.10E−2
	$h=1/10$	1.64E−1	3.75E−1	3.03E−2	2.36E−2	4.58E−2	2.05E−2
	$h=1/20$	8.26E−2	1.94E−1	1.52E−2	1.19E−2	2.29E−2	1.02E−2
	$h=1/40$	4.13E−2	9.79E−2	7.61E−3	6.02E−3	1.22E−2	5.13E−3
	$h=1/80$	2.06E−2	4.90E−2	3.80E−3	3.05E−3	6.68E−3	2.56E−3
	$h=1/160$	1.03E−2	2.45E−2	1.90E−3	1.56E−3	4.02E−3	1.28E−3

From Fig. 1 we can see the almost linear convergence of the iterative algorithm for lowest order Raviart-Thomas element with triangular and rectangular triangulation. The convergence results verify the effectiveness of the iterative algorithm.

6.3 Convergence of Mixed Element Approximation

The aim of this section is to show the convergence of the mixed element approximation. We test Examples 1 to 4 to get the convergence rate of the finite element approximation. 6 levels are computed to get the convergence rates, i.e. the discretization parameter $h = 1/5, 1/10, 1/20, 1/40, 1/80, 1/160$ respectively. The error control of the iterative algorithm is $eps = 10^{-4}$. The Forchheimer number is set to be 1, i.e. $\beta = 1$.

We use $\frac{\|\boldsymbol{u}-\boldsymbol{u}_h\|}{\|\boldsymbol{u}\|}$, $\frac{\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\|}{\|\nabla\cdot\boldsymbol{u}\|}$, $\frac{\|p-p_h\|}{\|p\|}$ as the criterion of the convergence of the finite element approximation. Here, $\|\cdot\|$ may be $\|\cdot\|_{0,2,\Omega}$, $\|\cdot\|_{0,3,\Omega}$, $\|\cdot\|_{0,\infty,\Omega}$.

Table 3 Relative error for rectangular triangulation

	Resolution	$E_{0,2}^{p,h}$	$E_{0,\infty}^{p,h}$	$E_{0,2}^{\boldsymbol{u},h}$	$E_{0,3}^{\boldsymbol{u},h}$	$E_{0,\infty}^{\boldsymbol{u},h}$	$E_{\mathrm{div},2}^{\boldsymbol{u},h}$
Ex1	$h=1/5$	2.66E−1	4.20E−1	6.96E−2	5.29E−2	1.26E−1	–
	$h=1/10$	1.30E−1	1.99E−1	3.48E−2	2.62E−2	6.51E−2	–
	$h=1/20$	6.46E−2	9.76E−2	1.74E−2	1.49E−2	3.32E−2	–
	$h=1/40$	3.22E−2	4.81E−2	8.70E−3	7.72E−3	1.69E−2	–
	$h=1/80$	1.61E−2	2.38E−2	4.36E−3	3.84E−3	9.01E−3	–
	$h=1/160$	8.08E−3	1.19E−2	2.21E−3	1.92E−3	5.11E−3	–
Ex2	$h=1/5$	3.91E−1	5.64E−1	6.96E−2	5.29E−2	1.26E−1	–
	$h=1/10$	2.01E−1	2.84E−1	3.48E−2	2.62E−2	6.51E−2	–
	$h=1/20$	1.01E−1	1.47E−1	1.74E−2	1.49E−2	3.32E−2	–
	$h=1/40$	5.06E−2	7.43E−2	8.70E−3	7.72E−3	1.69E−2	–
	$h=1/80$	2.53E−2	3.72E−2	4.36E−3	3.84E−3	9.01E−3	–
	$h=1/160$	1.26E−2	1.86E−2	2.21E−3	1.62E−3	5.11E−3	–
Ex3	$h=1/5$	2.75E−1	4.61E−1	5.76E−2	2.77E−2	8.74E−2	4.15E−2
	$h=1/10$	1.31E−1	1.96E−1	2.88E−2	1.43E−2	4.54E−2	2.08E−2
	$h=1/20$	6.48E−2	9.49E−2	1.44E−2	7.28E−3	2.29E−2	1.04E−2
	$h=1/40$	3.23E−2	4.72E−2	7.21E−3	3.60E−3	1.21E−2	5.20E−3
	$h=1/80$	1.61E−2	2.36E−2	3.61E−3	1.88E−3	6.56E−3	2.60E−3
	$h=1/160$	8.07E−3	1.18E−2	1.80E−3	9.42E−4	3.89E−3	1.30E−3
Ex4	$h=1/5$	3.91E−1	5.64E−1	5.76E−2	2.77E−2	8.74E−2	4.15E−2
	$h=1/10$	2.01E−1	2.85E−1	2.88E−2	1.43E−2	4.54E−2	2.08E−2
	$h=1/20$	1.01E−1	1.47E−1	1.44E−2	7.28E−3	2.29E−2	1.04E−2
	$h=1/40$	5.06E−2	7.43E−2	7.21E−3	3.60E−3	1.21E−2	5.20E−3
	$h=1/80$	2.53E−2	3.72E−2	3.61E−3	1.88E−3	6.56E−3	2.60E−3
	$h=1/160$	1.26E−2	1.86E−2	1.80E−3	9.42E−4	3.89E−3	1.30E−3

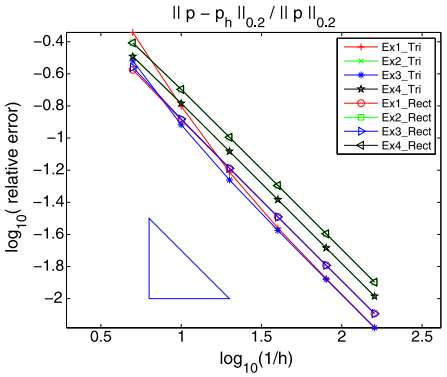
On the table we use the following symbols,

$$E_{0,2}^{p,h}=\frac{\|p-p_h\|_{0,2}}{\|p\|_{0,2}},\qquad E_{0,\infty}^{p,h}=\frac{\|p-p_h\|_{0,\infty}}{\|p\|_{0,\infty}},$$

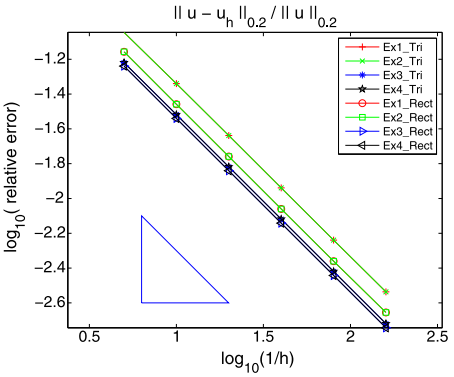
$$E_{0,2}^{\boldsymbol{u},h}=\frac{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,2}}{\|\boldsymbol{u}\|_{0,2}},\qquad E_{0,3}^{\boldsymbol{u},h}=\frac{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,3}}{\|\boldsymbol{u}\|_{0,3}},\qquad E_{0,\infty}^{\boldsymbol{u},h}=\frac{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\infty}}{\|\boldsymbol{u}\|_{0,\infty}},$$

$$E_{div,2}^{\boldsymbol{u},h}=\frac{\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\|_{0,2}}{\|\nabla\cdot\boldsymbol{u}\|_{0,2}}.$$

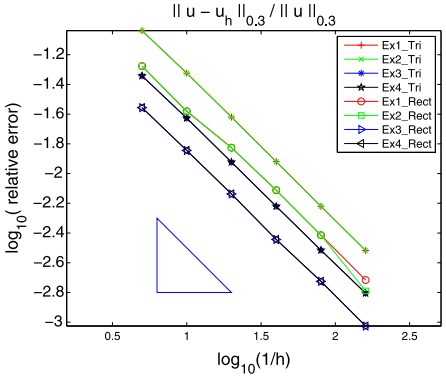
Tables 2 and 3 are the relative errors of Examples 1 to 4. Figure 2 displays the convergence rate in log-log scale. It is easy to see the convergence rate is almost first order.



(a) Convergence Rate for Pressure(L^2-norm)



(b) Convergence Rate for Velocity(L^2-norm)



(c) Convergence Rate for Velocity(L^3-norm)

Fig. 2 Convergence rate of mixed element. The tangent of the triangle is 1

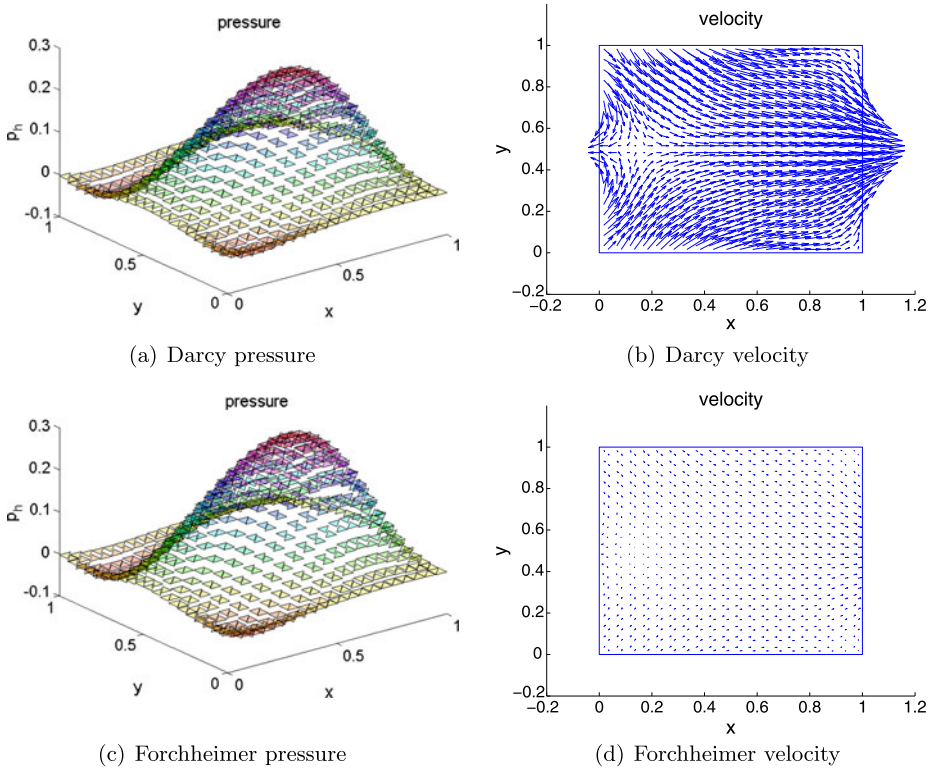
6.4 Comparison Between Forchheimer Flow and Darcy Flow

The purpose of this section is to observe the different behaviors of Darcy flow and Forchheimer flow. The discretization parameter $h = 1/20$. The error control of the iterative algorithm is $eps = 10^{-4}$. The Forchheimer number is set to be 100, i.e. $\beta = 100$.

We do experiment in different ways. Example 5 and 6 are without source, whose results are listed in Figs. 3 and 4, and Tables 4 and 5; Examples 7 and 8 are with source, whose results are listed in Figs. 5 and 6, and Tables 6 and 7. Here triangular element are used to do numerical examples.

Different behaviors of Darcy and Forchheimer flow are observed from Example 5 and 6, the problem without source. It is easy to see that the pressure distribution between Darcy and Forchheimer flow are almost the same, while the velocity field for Forchheimer flow is smaller than the Darcy flow.

Different behaviors of Darcy and Forchheimer flow are also observed from Examples 7 and 8. It is easy to see that the pressure distribution of Forchheimer flow is larger than the Darcy flow, while the velocity field for Forchheimer flow is smaller than the Darcy flow, which is similar to the cases without source.



(a) Darcy pressure (b) Darcy velocity

(c) Forchheimer pressure (d) Forchheimer velocity

Fig. 3 Example 5, triangular triangulation

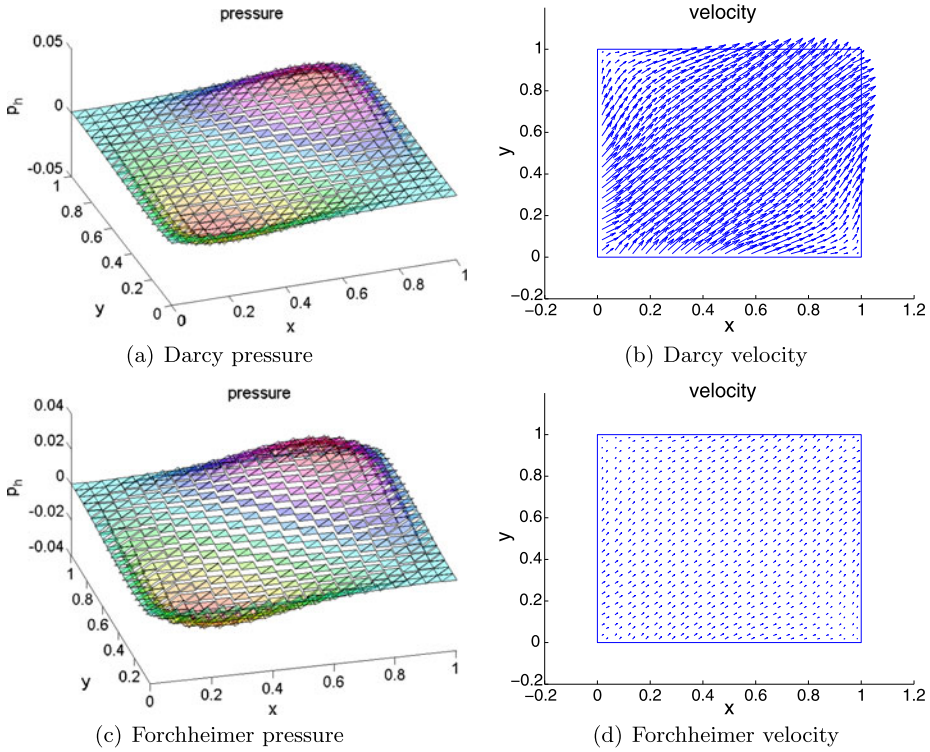
Table 4 Results for Example 5, $h = 1/20$

	$\|p_h\|_{0,2}$	$\|p_h\|_{0,\infty}$	$\|\boldsymbol{u}_h\|_{0,2}$	$\|\boldsymbol{u}_h\|_{0,3}$	$\|\boldsymbol{u}_h\|_{0,\infty}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,2}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,\infty}$
Darcy	0.0984	0.24	0.8184	0.7892	2.1859	0	0
Forchheimer	0.1134	0.2745	0.0829	0.0788	0.1807	0	0

Table 5 Results for Example 6, $h = 1/20$

	$\|p_h\|_{0,2}$	$\|p_h\|_{0,\infty}$	$\|\boldsymbol{u}_h\|_{0,2}$	$\|\boldsymbol{u}_h\|_{0,3}$	$\|\boldsymbol{u}_h\|_{0,\infty}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,2}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,\infty}$
Darcy	0.0118	0.0241	0.2426	0.2206	0.4417	0	0
Forchheimer	0.0127	0.0254	0.0448	0.0405	0.0756	0	0

From these numerical examples we can see that, for the problem with source, the Forchheimer term causes the pressure to rise; for the problem without source, the Forchheimer term doesn't affect the pressure distribution. In both case the velocity field of the Forchheimer flow is smaller than that of Darcy flow.



(a) Darcy pressure (b) Darcy velocity

(c) Forchheimer pressure (d) Forchheimer velocity

Fig. 4 Example 6, triangular triangulation

Table 6 Results for Example 7, $h = 1/20$

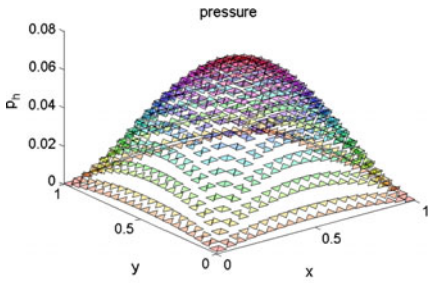
	$\|p_h\|_{0,2}$	$\|p_h\|_{0,\infty}$	$\|\boldsymbol{u}_h\|_{0,2}$	$\|\boldsymbol{u}_h\|_{0,3}$	$\|\boldsymbol{u}_h\|_{0,\infty}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,2}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,\infty}$
Darcy	0.0376	0.0683	1.0143	1.0142	1.315	0.8556	0.9992
Forchheimer	0.7758	1.3029	0.175	0.1745	0.3342	0.8556	0.9992

Table 7 Results for Example 8, $h = 1/20$

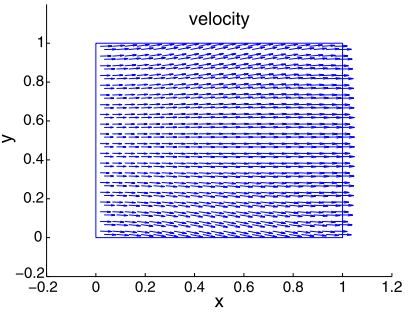
	$\|p_h\|_{0,2}$	$\|p_h\|_{0,\infty}$	$\|\boldsymbol{u}_h\|_{0,2}$	$\|\boldsymbol{u}_h\|_{0,3}$	$\|\boldsymbol{u}_h\|_{0,\infty}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,2}$	$\|\nabla\cdot\boldsymbol{u}_h\|_{0,\infty}$
Darcy	0.0101	0.0201	1.0025	1.0010	1.1315	0.4983	0.9898
Forchheimer	0.1435	0.2849	0.1093	0.0990	0.2097	0.4983	0.9898

Acknowledgements The work is supported by the National Natural Science Foundation of China Grant No. 11171190.

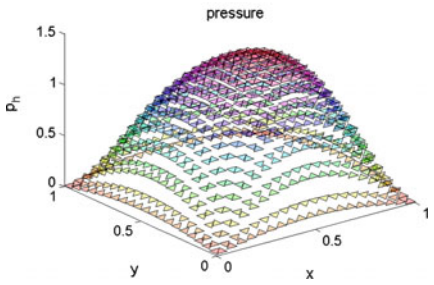
The authors thank the two anonymous referees for their constructive corrections and comments, which lead to improvements of the presentation.



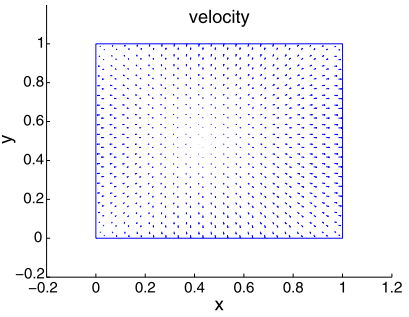
(a) Darcy pressure



(b) Darcy velocity

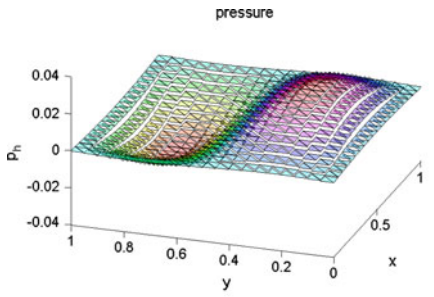


(c) Forchheimer pressure

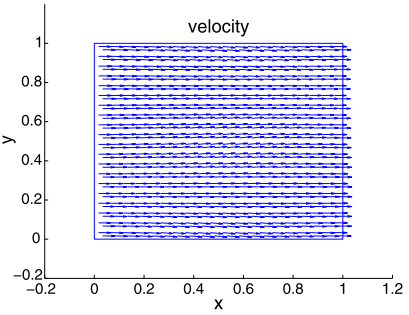


(d) Forchheimer velocity

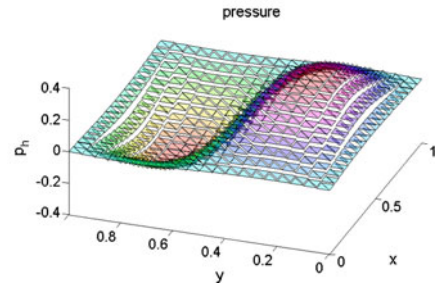
Fig. 5 Example 7, triangular triangulation



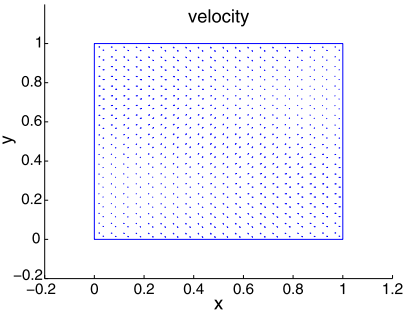
(a) Darcy pressure



(b) Darcy velocity



(c) Forchheimer pressure



(d) Forchheimer velocity

Fig. 6 Example 8, triangular triangulation

References

1. Aziz, K., Settari, A.: Petroleum Reservoir Simulation. Applied Science Publishers LTD, London (1979)
2. Bahriawati, C., Carstensen, C.: Three Matlab implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control. Comput. Methods Appl. Math. **5**(4), 333–361 (2005)
3. Bergamaschi, L., Mantica, S., Saleri, F.: Mixed finite element approximation of Darcy's law in porous media. Technical report CRS4-AppMath-94-17, CRS4 (1994)
4. Brezzi, F., Douglas, J. Jr., Duran, R., Fortin, M.: Mixed finite elements for second order elliptic problems in three variables. Numer. Math. **51**(2), 237–250 (1987)
5. Brezzi, F., Douglas, J. Jr., Fortin, M., Marini, L.D.: Efficient rectangular mixed finite elements in two and three space variables. Math. Model. Numer. Anal. **21**(4), 581–604 (1987)
6. Brezzi, F., Douglas, J. Jr., Marini, L.D.: Two families of mixed finite elements for second order elliptic problems. Numer. Math. **47**(2), 217–235 (1985)
7. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer Series in Computational Mathematics, vol. 15. Springer, Berlin (1991)
8. Chen, Z.: Finite Element Methods and Their Applications. Scientific Computation. Springer, Berlin (2005)
9. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Studies in Mathematics and Its Applications, vol. 4. North-Holland, Amsterdam (1978)
10. Farhloul, M.: A mixed finite element method for a nonlinear Dirichlet problem. IMA J. Numer. Anal. **18**, 121–132 (1998)
11. Farhloul, M., Manouzi, H.: On a mixed finite element method for the p-Laplacian. Can. Appl. Math. Q. **8**(1), 67–78 (2000)
12. Girault, V., Raviart, P.-A.: Finite Element Methods for Navier-Stokes Equation: Theory and Algorithm. Springer Series in Computational Mathematics, vol. 5. Springer, Berlin (1986)
13. Girault, V., Wheeler, M.F.: Numerical discretization of a Darcy-Forchheimer model. Numer. Math. **110**(2), 161–198 (2008)
14. Glowinski, R., Marroco, A.: Sur lapproximation, par lments finis dordre un, et la rsolution, par pnalisation-dualit dune classe de problmes de Dirichlet non linaires. RAIRO. Anal. Numér. **9**(2), 41–76 (1975)
15. Lopez, H., Molina, B., Jose, J.S.: Comparison between different numerical discretizations for a Darcy-Forchheimer model. Electron. Trans. Numer. Anal. **34**, 187–203 (2009)
16. Nedelec, J.C.: Mixed finite elements in R3. Numer. Math. **35**(3), 315–341 (1980)
17. Nedelec, J.C.: A new family of mixed finite elements in R3. Numer. Math. **50**(1), 57–81 (1986)
18. Neuman, S.P.: Theoretical derivation of Darcy's law. Acta Mech. **25**(3), 153–170 (1977)
19. Raviart, P.-A., Thomas, J.M.: A mixed finite element method for 2-nd order elliptic problems. In: Mathematical Aspects of the Finite Element Method. Lecture Notes in Mathematics, vol. 606, pp. 292–315. Springer, Berlin (1977). Chap. 19
20. Roberts, J.E., Thomas, J.-M.: Mixed and hybrid methods. In: Finite Element Methods. Part 1. Handbook of Numerical Analysis, vol. 2, pp. 523–639. Elsevier Science Publishers B.V., Amsterdam (1991)
21. Ruth, D., Ma, H.: On the derivation of the Forchheimer equation by means of the averaging theorem. Transp. Porous Media **7**(3), 255–264 (1992)
22. Sandri, D.: Sur lapproximation numérique des coulements quasi-newtoniens dont la viscosit suit la loi puissance ou la loi de Carreau. RAIRO. Modél. Math. Anal. Numér. **27**(2), 131–155 (1993)
23. Urquiza, J.M., N' Dri, D., Garon, A., Delfour, M.C.: A numerical study of primal mixed finite element approximations of Darcy equations. Commun. Numer. Methods Eng. **22**, 901–915 (2006)
24. Whitaker, S.: Flow in porous media I: A theoretical derivation of Darcy's law. Transp. Porous Media **1**(1), 3–25 (1986)