Journal of Computational Mathematics
Vol.xx, No.x, 2024, 1–25.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2404-m2023-0232

# WEAK GALERKIN METHOD FOR COUPLING STOKES AND DARCY-FORCHHEIMER FLOWS*

Mingze Qin
*School of Mathematics, Jilin University, Changchun 130012, China*
*Email: qinmz21@mail.jlu.edu.cn*
Hui Peng
*School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, China*
*Email: penghui23@sjtu.edu.cn*
Qilong Zhai[1)]
*School of Mathematics, Jilin University, Changchun 130012, China*
*Email: zhaiql@jlu.edu.cn*

**Abstract**

In this paper, we introduce the weak Galerkin (WG) method for solving the coupled Stokes and Darcy-Forchheimer flows problem with the Beavers-Joseph-Saffman interface condition in bounded domains. We define the WG spaces in the polygonal meshes and construct corresponding discrete schemes. We prove the existence and uniqueness of the WG scheme by the discrete inf-sup condition and monotone operator theory. Then, we derive the optimal error estimates for the velocity and pressure. Numerical experiments are presented to verify the efficiency of the WG method.

*Mathematics subject classification:* 65N30, 65N15, 76S05, 65N12.
*Key words:* Weak Galerkin method, Coupled Stokes and Darcy-Forchheimer flows, Monotone operator.

## 1. Introduction

In recent years, the coupled model of free fluid and porous medium fluid has received more and more attention. Such problems have been widely applied in groundwater flows [9, 13, 17], environmental science [15], flow in vuggy porous media [2, 4] and so on. The classical coupled Stokes-Darcy model consists of the Stokes equation in fluid region, the Darcy's law in porous medium region. The interface conditions in the model are the flow continuity condition, the force equilibrium condition, and the Beavers-Joseph-Saffman condition [34]. However, Darcy's law describes the porous media flow at low speed. In the case of high-speed seepage, porous media will be turbulent. Forchheimer found through experiments that the pressure gradient and the Darcy velocity should meet the nonlinear relationship with a coefficient depending on the soil grain diameter and porosity for the larger values of Reynolds number, which requires the use of the modified Darcy equation in 1901, namely the Forchheimer equation. The Darcy-Forchheimer model in porous media explains the nonlinear behavior by adding an item representing the inertial effect, which is used in the exploration and production of oil and gas in petroleum reservoir [6].

* Received October 17, 2023 / Revised version received December 29, 2023 / Accepted April 16, 2024 / Published online June 13, 2024 /
1) Corresponding author

So far, there are many types of research on the coupling of Stokes flow and Darcy flow, various numerical methods have been developed for the Stokes-Darcy model [3, 10, 14, 16, 18, 20, 26, 28, 29, 32], and there is a lot of research on the individual Darcy-Forchheimer model [5, 25, 30, 31, 33]. However, there are relatively few researches on the coupled Stokes-Darcy-Forchheimer model. Zhang and Rui [42] used the Crouzeix-Raviart element to approximate the velocity and obtained the optimal-order error estimates when the velocity and pressure were in $[H^2]^N$ ($N = 2$ or 3) and $H^1$ spaces respectively. The authors also derived a discrete inf-sup condition and establish the existence and uniqueness of the problem. Zhao *et al.* [45] proposed the staggered discontinuous Galerkin method for the model problems, which could be flexibly applied to rough grids such as the highly distorted grids and the polygonal grids. And they proved optimal convergence estimates by proposing some new discrete trace inequality and generalized Poincaré-Friedrichs inequality. The fully mixed finite element method was proposed by A. Almonacid *et al.* [1]. They considered the Darcy-Forchheimer model by means of a modified abstract theory for twofold saddle point problems, which could pose the variational formulation in terms of just Banach spaces, and the optimal-order error estimates were obtained.

The goal of this paper is to investigate the weak Galerkin (WG) method for solving the coupled Stokes and Darcy-Forchheimer flows problem. The WG method is first proposed for the second-order elliptic equations by Wang and Ye [36], and further developed in [23,35,37,39,43]. The key idea of the WG method is to use discontinuous piecewise polynomials as basis functions, and to replace the classical derivative operators by specifically defined weak derivative operators in the numerical scheme. In the past few years, the WG methods have been widely applied to [10,21,22,38,40,41,44]. The applications mentioned above are mostly linear problems and the theoretical analysis and numerical experiments are relatively more complicated for nonlinear problems.

In this paper, we first establish the WG numerical scheme on polygonal meshes. Due to the nonlinear and monotonicity of the Forchheimer term, we use Minty-Browder theorem and discrete inf-sup condition to prove the existence and uniqueness of the numerical scheme. Furthermore, by using the technical estimate for the Forchheimer term, we obtain the optimal orders of error estimates when the velocity and pressure are in the $[H^{k+1}]^2$ space and the $H^k$ space, respectively. When conducting numerical experiments, we adopted triangular grids, and the numerical results are consistent with the theoretical estimates.

This paper is organized as follows. In Section 2, we introduce the model problem and some notations. In Section 3, we define some WG spaces and establish the numerical scheme. In Section 4, we prove the existence and uniqueness of the WG scheme. The error estimates for the WG approximations are given in Section 5. Some numerical experiments are presented in Section 6.

## 2. Model Problem

In this section, we describe in detail the coupled Stokes and Darcy-Forchheimer model [42]. We consider the coupled flow in a bounded domain $\Omega \in \mathbb{R}^2$, which consists of two subregions, a free region $\Omega_s$ and a porous medium region $\Omega_d = \Omega\backslash\overline{\Omega}_s$. Both $\Omega_s$ and $\Omega_d$ have Lipschitz continuous boundaries. Define $\Gamma_I = \partial\Omega_s \cap \partial\Omega_d, \Gamma_s = \partial\Omega_s\backslash\Gamma_I, \Gamma_d = \partial\Omega_d\backslash\Gamma_I$. $\boldsymbol{n}_s$ and $\boldsymbol{n}_d$ are the outward unit normal vectors on $\Omega_s$ and $\Omega_d, \boldsymbol{\tau}$ is the unit tangent vector on $\Gamma_I$, as shown in Fig. 2.1.

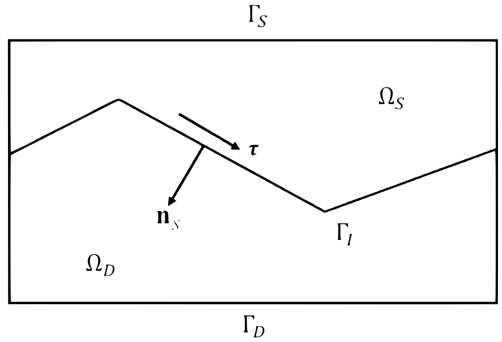


Fig. 2.1. Domain of the coupled Stokes and Darcy-Forchheimer Flows.

In the whole domain $\Omega$, we denote the fluid velocity by $\boldsymbol{u}$ and the pressure by $p$. In the free flow region $\Omega_s$, the fluid flow is governed by the Stokes equations. The velocity $\boldsymbol{u}_s$ and the pressure $p_s$ satisfy

$$-2\mu\nabla\cdot\boldsymbol{D}(\boldsymbol{u}_s)+\nabla p_s=\boldsymbol{f}_s \quad \text{in } \Omega_s, \tag{2.1}$$

$$\nabla\cdot\boldsymbol{u}_s=0 \quad \text{in } \Omega_s, \tag{2.2}$$

$$\boldsymbol{u}_s=\mathbf{0} \quad \text{on } \Gamma_s. \tag{2.3}$$

In the porous medium region $\Omega_d$, we use Darcy-Forchheimer equations to describe the flow motion. Find velocity $\boldsymbol{u}_d$ and pressure $p_d$ such that

$$\mu\mathbf{K}^{-1}\boldsymbol{u}_d+\beta\rho|\boldsymbol{u}_d|\boldsymbol{u}_d+\nabla p_d=\boldsymbol{f}_d \quad \text{in } \Omega_d, \tag{2.4}$$

$$\nabla\cdot\boldsymbol{u}_d=g_d \quad \text{in } \Omega_d, \tag{2.5}$$

$$\boldsymbol{u}_d\cdot\boldsymbol{n}_d=0 \quad \text{on } \Gamma_d. \tag{2.6}$$

On the interface $\Gamma_I$, we impose the following interface conditions:

$$\boldsymbol{u}_s\cdot\boldsymbol{n}_s+\boldsymbol{u}_d\cdot\boldsymbol{n}_d=0, \tag{2.7}$$

$$p_s-2\mu\boldsymbol{n}_s\cdot\boldsymbol{D}(\boldsymbol{u}_s)\cdot\boldsymbol{n}_s=p_d, \tag{2.8}$$

$$-\frac{\sqrt{\kappa}}{\alpha}2\boldsymbol{n}_s\cdot\boldsymbol{D}(\boldsymbol{u}_s)\cdot\boldsymbol{\tau}=\boldsymbol{u}_s\cdot\boldsymbol{\tau}, \tag{2.9}$$

where (2.9) is the Beavers-Joseph-Saffman condition [34].

Here, $\mu>0$ is the fluid viscosity, $\beta>0$ is the dynamic viscosity, $\rho>0$ is the density of the fluid, $\boldsymbol{f}_s$, $\boldsymbol{f}_d$ and $g_d$ are in $[L^2(\Omega_s)]^2$, $[L^2(\Omega_d)]^2$ and $L^2(\Omega_d)$, respectively, and $\boldsymbol{D}$ is the symmetric strain tensor defined by

$$\boldsymbol{D}(\boldsymbol{u})=\frac{1}{2}(\nabla\boldsymbol{u}+\nabla\boldsymbol{u}^\top).$$

Besides, $\mathbf{K}$ is the symmetric positive define tensor, which is uniformly bounded and smooth in $\Omega_d$. In (2.9), $\kappa=\boldsymbol{\tau}\cdot\mathbf{K}\cdot\boldsymbol{\tau}$, and $\alpha$ is a parameter determined by the experiment.

## 3. The Weak Galerkin Scheme

In this section, we establish the WG scheme for model problem (2.1)-(2.9). Throughout the paper, we follow the usual notations for Sobolev spaces $W^{m,p}(\Omega)$ and norms $\|\cdot\|_{W^{m,p}(\Omega)}$. The space $H(\text{div};\Omega)$ is defined by

$$H(\mathrm{div};\Omega)=\{\boldsymbol{v}:\boldsymbol{v}\in[L^2(\Omega)]^2,\nabla\cdot\boldsymbol{v}\in L^2(\Omega)\}.$$

Let $\mathcal{T}_h$ be a polygonal partition of $\Omega$ satisfying shape regularity assumptions described in [37]. $\mathcal{T}_{s,h}$ and $\mathcal{T}_{d,h}$ are the restriction of $\mathcal{T}_h$ in $\Omega_s$ and $\Omega_d$, respectively. Denote by $T_l$ $(l=s,d)$ a cell in $\mathcal{T}_{l,h}$ and by $E_l$ an edge in $\mathcal{T}_{l,h}$. For each cell $T_l\in\mathcal{T}_{l,h}, h_{T_l}$ is its diameter, and $h=\max_{T\in\mathcal{T}_h}h_T$. Let $\mathcal{E}_{l,h}$ and $\mathcal{E}_{I,h}$ be the set of all the edges in $\mathcal{T}_h\cap(\Omega_l\cup\Gamma_l)$ and $\mathcal{T}_h\cap\Gamma_I$, respectively. $\mathcal{E}_{l,h}^0=\mathcal{E}_{l,h}\setminus\partial\Omega$ be the set of all interior edges in $\mathcal{T}_h$, $\mathcal{E}_{l,h}^b=\mathcal{E}_{l,h}\cap\Gamma_l$ be the set of all exterior edges in $\mathcal{T}_h$.

Let $P_k(T)$ be the space of all polynomials in $T$ with degree no more than $k$. For a given integer $k\geqslant 1$, we define the weak Galerkin finite element spaces as follows:

$$\begin{aligned}
V_{s,h}=&\big\{\boldsymbol{v}_{s,h}=\{\boldsymbol{v}_{s,0},\boldsymbol{v}_{s,b}\},\ \boldsymbol{v}_{s,0|_{T_s}}\in[P_k(T_s)]^2,\ T_s\in\mathcal{T}_{s,h},\\
&\quad \boldsymbol{v}_{s,b|_{E_s}}\in[P_k(E_s)]^2,\ E_s\in\mathcal{E}_{s,h}\cup\mathcal{E}_{I,h},\ \boldsymbol{v}_{s,b}=0\text{ on }\mathcal{E}_{s,h}^b\big\},\\
V_{d,h}=&\big\{\boldsymbol{v}_{d,h}=\{\boldsymbol{v}_{d,0},\boldsymbol{v}_{d,b}\},\ \boldsymbol{v}_{d,0|_{T_d}}\in[P_k(T_d)]^2,\ T_d\in\mathcal{T}_{d,h},\\
&\quad \boldsymbol{v}_{d,b|_{E_d}}=v_{d,b}\boldsymbol{n_d},\ v_{d,b}\in P_k(E_d),\ E_d\in\mathcal{E}_{d,h},\ \boldsymbol{v}_{d,b}=0\text{ on }\mathcal{E}_{d,h}^b\big\},\\
V_h=&\{\boldsymbol{v}_h=(\boldsymbol{v}_{s,h},\boldsymbol{v}_{d,h})\in V_{s,h}\times V_{d,h}\},\\
W_{s,h}=&\big\{p_{s,h}:p_{s,h}\in L^2(\Omega_s),\ p_{s,h|_{T_s}}\in P_{k-1}(T_s),\ \forall\,T_s\in\mathcal{T}_{s,h}\big\},\\
W_{d,h}=&\big\{p_{d,h}:p_{d,h}\in L^2(\Omega_d),\ p_{d,h|_{T_d}}\in P_{k-1}(T_d),\ \forall\,T_d\in\mathcal{T}_{d,h}\big\},\\
W_h=&\{p_h=(p_{s,h},p_{d,h})\in W_{s,h}\times W_{d,h};\ (p_{s,h},1)_{\Omega_s}+(p_{d,h},1)_{\Omega_d}=0\}.
\end{aligned}$$

Now we introduce some projection operators. For each cell $T_l$, let $Q_{l,0}$ be the $L^2$ projection operator onto $[P_k(T_l)]^2$, $\mathbb{Q}_{l,h}$ be the $L^2$ projection operator onto $[P_{k-1}(T_l)]^{2\times 2}$, and $\pi_{l,h}$ be the $L^2$ projection operator onto $P_{k-1}(T_l)$. On each edge $E_l$, denoted by $Q_{l,b}$ the $L^2$ projection operator onto $[P_k(E_l)]^2$. Combining $Q_{l,0}$ and $Q_{l,b}$, we define $Q_{l,h}=\{Q_{l,0},Q_{l,b}\}$ the projection operator onto $V_{l,h}$. Then, define $Q_h=(Q_{s,h},Q_{d,h}),\mathbb{Q}_h=(\mathbb{Q}_{s,h},\mathbb{Q}_{d,h})$ and $\pi_h=(\pi_{s,h},\pi_{d,h})$. After that, we denote $(v,w)_{T_l}$ as $\int_{T_l}vw\mathrm{d}x$ and $\langle v,w\rangle_{\partial T_l}$ as $\int_{\partial T_l}vw\mathrm{d}s$.

In the weak Galerkin finite element spaces $V_h$, we define the weak gradient and weak divergence operators as follows.

**Definition 3.1 ([37]).** *For any $\boldsymbol{v}_h=\{\boldsymbol{v}_0,\boldsymbol{v}_b\}\in V_h$, on each cell $T\in\mathcal{T}_h$, define $\nabla_w\boldsymbol{v}_h$ and $\nabla_w\cdot\boldsymbol{v}_h$ by*

$$\begin{aligned}
(\nabla_w\boldsymbol{v}_h,\gamma)_T&=-(\boldsymbol{v}_0,\nabla\cdot\gamma)_T+\langle\boldsymbol{v}_b,\gamma\boldsymbol{n}\rangle_{\partial T},\quad \forall\,\gamma\in[P_{k-1}(T)]^{2\times 2},\\
(\nabla_w\cdot\boldsymbol{v}_h,q)_T&=-(\boldsymbol{v}_0,\nabla q)_T+\langle\boldsymbol{v}_b\cdot\boldsymbol{n},q\rangle_{\partial T},\quad \forall\,q\in P_{k-1}(T),
\end{aligned}$$

*where $\boldsymbol{n}$ is the unit outward normal vector of $\partial T$.*

Define

$$\boldsymbol{D}_w(\boldsymbol{u})=\frac{1}{2}\big(\nabla_w\boldsymbol{u}+\nabla_w\boldsymbol{u}^\top\big).$$

Let $\boldsymbol{u}_h=(\boldsymbol{u}_{s,h},\boldsymbol{u}_{d,h}),\boldsymbol{v}_h=(\boldsymbol{v}_{s,h},\boldsymbol{v}_{d,h})\in V_h$ and $p_h=(p_{s,h},p_{d,h})\in W_h$. Now we introduce the following five bilinear forms:

$$\begin{aligned}
a_s(\boldsymbol{u}_{s,h},\boldsymbol{v}_{s,h})&=2\mu\sum_{T_s\in\mathcal{T}_{s,h}}\big(\boldsymbol{D}_w(\boldsymbol{u}_{s,h}),\boldsymbol{D}_w(\boldsymbol{v}_{s,h})\big)_{T_s}+\frac{\mu\alpha}{\sqrt{\kappa}}\langle\boldsymbol{u}_{s,b}\cdot\boldsymbol{\tau},\boldsymbol{v}_{s,b}\cdot\boldsymbol{\tau}\rangle_{\Gamma_I},\\
b_s(\boldsymbol{v}_{s,h},p_{s,h})&=(p_{s,h},\nabla_w\cdot\boldsymbol{v}_{s,h})_{\Omega_s},
\end{aligned}$$

$$
\begin{aligned}
b_d(\boldsymbol{v}_{d,h}, p_{d,h}) &= (p_{d,h}, \nabla_w \cdot \boldsymbol{v}_{d,h})_{\Omega_d},\\
s_s(\boldsymbol{u}_{s,h}, \boldsymbol{v}_{s,h}) &= \rho_s \sum_{T_s \in \mathcal{T}_{s,h}} h_{T_s}^{-1} \langle \boldsymbol{u}_{s,0} - \boldsymbol{u}_{s,b}, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s},\\
s_d(\boldsymbol{u}_{d,h}, \boldsymbol{v}_{d,h}) &= \rho_d \sum_{T_d \in \mathcal{T}_{d,h}} h_{T_d}^{-1} \langle (\boldsymbol{u}_{d,0} - \boldsymbol{u}_{d,b}) \cdot \boldsymbol{n}_d, (\boldsymbol{v}_{d,0} - \boldsymbol{v}_{d,b}) \cdot \boldsymbol{n}_d \rangle_{\partial T_d},
\end{aligned}
$$

where $\rho_s$ and $\rho_d$ can be any positive number. And we define a nonlinear form on $V_{d,h}$

$$
a_d(\boldsymbol{u}_{d,h}, \boldsymbol{v}_{d,h}) = \mu\big(\boldsymbol{K}^{-1}\boldsymbol{u}_{d,0}, \boldsymbol{v}_{d,0}\big)_{\Omega_d} + (\beta\rho|\boldsymbol{u}_{d,0}|\boldsymbol{u}_{d,0}, \boldsymbol{v}_{d,0})_{\Omega_d}.
$$

Let

$$
\begin{aligned}
a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) &= a_s(\boldsymbol{u}_{s,h}, \boldsymbol{v}_{s,h}) + a_d(\boldsymbol{u}_{d,h}, \boldsymbol{v}_{d,h}),\\
b_h(\boldsymbol{v}_h, p_h) &= b_s(\boldsymbol{v}_{s,h}, p_{s,h}) + b_d(\boldsymbol{v}_{d,h}, p_{d,h}),\\
s(\boldsymbol{u}_h, \boldsymbol{v}_h) &= s_s(\boldsymbol{u}_{s,h}, \boldsymbol{v}_{s,h}) + s_d(\boldsymbol{u}_{d,h}, \boldsymbol{v}_{d,h}).
\end{aligned}
$$

With these preparations, we introduce the following WG scheme for the problem (2.1)-(2.9).

**Algorithm 3.1:** Weak Galerkin Algorithm.

Find $(\boldsymbol{u}_h, p_h) \in V_h \times W_h$ such that

$$
\begin{cases}
a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) - b_h(\boldsymbol{v}_h, p_h) + s(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_0), & \forall\, \boldsymbol{v}_h \in V_h,\\
b_h(\boldsymbol{u}_h, q_h) = (g, q_h), & \forall\, q_h \in W_h,
\end{cases}
\tag{3.1}
$$

where $(\boldsymbol{f}, \boldsymbol{v}_0) = (\boldsymbol{f}_s, \boldsymbol{v}_{s,0})_{\Omega_s} + (\boldsymbol{f}_d, \boldsymbol{v}_{d,0})_{\Omega_d}$ and $(g, q_h) = (0, q_{s,h})_{\Omega_s} + (g_d, q_{d,h})_{\Omega_d}$.

## 4. Existence and Uniqueness

In this section, we are concerned with the existence and uniqueness of the WG scheme (3.1). The key point is to verify the Minty-Browder theorem and the discrete inf-sup condition. First, we introduce discrete norms in $V_h$ and $W_h$. For any $\boldsymbol{v}_h \in V_h, p_h \in W_h$, define

$$
\begin{aligned}
|\!|\!| \boldsymbol{v}_h |\!|\!| = \Bigg( & 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \|\boldsymbol{D}_w(\boldsymbol{v}_{s,h})\|^2_{L^2(T_s)} + \frac{\mu\alpha}{\sqrt{\kappa}} \|\boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau}\|^2_{L^2(\Gamma_I)}\\
& + \rho_s \sum_{T_s \in \mathcal{T}_{s,h}} h_{T_s}^{-1} \|\boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b}\|^2_{L^2(\partial T_s)} + \rho_d \sum_{T_d \in \mathcal{T}_{d,h}} h_{T_d}^{-1} \|(\boldsymbol{v}_{d,0} - \boldsymbol{v}_{d,b}) \cdot \boldsymbol{n}_d\|^2_{L^2(\partial T_d)}\\
& + \mu \big\|\mathbf{K}^{-\frac{1}{2}} \boldsymbol{v}_{d,0}\big\|^2_{L^2(\Omega_d)} + \beta\rho \|\boldsymbol{v}_{d,0}\|^2_{L^3(\Omega_d)} + \|\nabla_w \cdot \boldsymbol{v}_h\|^2_{L^2(\Omega)} \Bigg)^{\frac{1}{2}},
\end{aligned}
$$

$$
\begin{aligned}
|\!|\!| \boldsymbol{v}_h |\!|\!|_1 &= \big( |\!|\!| \boldsymbol{v}_h |\!|\!|^2 - \beta\rho \|\boldsymbol{v}_{d,0}\|^2_{L^3(\Omega_d)} \big)^{\frac{1}{2}},\\
\|p_h\|_h &= \big( \|p_{s,h}\|^2_{L^2(\Omega_s)} + \|p_{d,h}\|^2_{L^2(\Omega_d)} \big)^{\frac{1}{2}}.
\end{aligned}
$$

It is easy to know that $|\!|\!| \cdot |\!|\!|_1$ is a norm in the $V_h$ and $\|\cdot\|_h$ is a norm in the $W_h$. The proof can be referred to [10, Section 3.1]. Clearly, $|\!|\!| \cdot |\!|\!|$ is a norm in the $V_h$, too.

Next, we give the inf-sup condition.

**Lemma 4.1.** *There exists a constant $\varrho > 0$ independent $h$ such that*

$$\inf_{p_h\in W_h}\sup_{\boldsymbol{v}_h\in V_h}\frac{b_h(\boldsymbol{v}_h,p_h)}{\|p_h\|_h\|\boldsymbol{v}_h\|}\geq\varrho. \tag{4.1}$$

*Proof.* Follow [7], for any $p_h \in W_h$, there exists $\widetilde{\boldsymbol{v}} = (\widetilde{\boldsymbol{v}}_s, \widetilde{\boldsymbol{v}}_d) \in [H_0^1(\Omega_s)]^2 \times [H_0^1(\Omega_d)]^2$ such that $\nabla\cdot\widetilde{\boldsymbol{v}}_s = p_{s,h}$ and $\nabla\cdot\widetilde{\boldsymbol{v}}_d = p_{d,h}$, satisfying

$$\|\widetilde{\boldsymbol{v}}_s\|_{H^1(\Omega_s)}+\|\widetilde{\boldsymbol{v}}_d\|_{H^1(\Omega_d)}\leq C(\|p_{s,h}\|_{\Omega_s}+\|p_{d,h}\|_{\Omega_d}), \tag{4.2}$$

where $C > 0$ is a constant only depending $\Omega$. By setting $\boldsymbol{v}_h = Q_h\widetilde{\boldsymbol{v}} \in V_h$, we claim that

$$\|Q_h\widetilde{\boldsymbol{v}}\|\leq C\big(\|\widetilde{\boldsymbol{v}}_s\|_{H^1(\Omega_s)}+\|\widetilde{\boldsymbol{v}}_d\|_{H^1(\Omega_d)}\big). \tag{4.3}$$

In fact, from Lemma A.1, we have

$$\begin{aligned}\sum_{T_s\in\mathcal{T}_{s,h}}\|\boldsymbol{D}_w(Q_{s,h}\widetilde{\boldsymbol{v}}_s)\|^2_{L^2(T_s)}&\leq C\sum_{T_s\in\mathcal{T}_{s,h}}\|\nabla_w(Q_{s,h}\widetilde{\boldsymbol{v}}_s)\|^2_{L^2(T_s)}\\&=C\sum_{T_s\in\mathcal{T}_{s,h}}\|\mathbb{Q}_{s,h}(\nabla\widetilde{\boldsymbol{v}}_s)\|^2_{L^2(T_s)}\leq C\|\nabla\widetilde{\boldsymbol{v}}_s\|^2_{L^2(\Omega_s)},\end{aligned} \tag{4.4}$$

and

$$\begin{aligned}\|\nabla_w\cdot(Q_h\widetilde{\boldsymbol{v}})\|^2_{L^2(\Omega)}&\leq\|\pi_h(\nabla\cdot\widetilde{\boldsymbol{v}})\|^2_{L^2(\Omega)}\leq\|\nabla\cdot\widetilde{\boldsymbol{v}}\|^2_{L^2(\Omega)}\\&\leq\|\nabla\widetilde{\boldsymbol{v}}\|^2_{L^2(\Omega_s)}+\|\nabla\widetilde{\boldsymbol{v}}\|^2_{L^2(\Omega_d)}.\end{aligned} \tag{4.5}$$

Using Lemma A.5, we obtain

$$\begin{aligned}\|Q_{s,b}\widetilde{\boldsymbol{v}}_s\cdot\boldsymbol{\tau}\|^2_{L^2(\Gamma_I)}&\leq C\sum_{E_s\in\mathcal{E}_{I,h}}\|Q_{s,b}\widetilde{\boldsymbol{v}}_s\|^2_{L^2(E_s)}\leq C\sum_{E_s\in\mathcal{E}_{I,h}}\|\widetilde{\boldsymbol{v}}_s\|^2_{L^2(E_s)}\\&\leq C\|\widetilde{\boldsymbol{v}}_s\|^2_{L^2(\partial\Omega_s)}\leq C\|\widetilde{\boldsymbol{v}}_s\|^2_{H^1(\Omega_s)}.\end{aligned} \tag{4.6}$$

According to the trace inequality (A.3), the projection inequality (A.5) and (A.6), it follows that

$$\begin{aligned}\rho_s\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\|Q_{s,0}\widetilde{\boldsymbol{v}}_s-Q_{s,b}\widetilde{\boldsymbol{v}}_s\|^2_{L^2(\partial T_s)}&\leq C\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\|Q_{s,0}\widetilde{\boldsymbol{v}}_s-\widetilde{\boldsymbol{v}}_s\|^2_{L^2(\partial T_s)}\\&\leq C\|\widetilde{\boldsymbol{v}}_s\|^2_{H^1(\Omega_s)}.\end{aligned} \tag{4.7}$$

Similarly,

$$\rho_d\sum_{T_d\in\mathcal{T}_{d,h}}h_{T_d}^{-1}\|(Q_{d,0}\widetilde{\boldsymbol{v}}_d-Q_{d,b}\widetilde{\boldsymbol{v}}_d)\cdot\boldsymbol{n}_d\|^2_{L^2(\partial T_d)}\leq C\|\widetilde{\boldsymbol{v}}_d\|^2_{H^1(\Omega_d)}. \tag{4.8}$$

From Lemma A.5 and the projection inequality (A.5), we have

$$\big\|\mathbf{K}^{-\frac{1}{2}}Q_{d,0}\widetilde{\boldsymbol{v}}_d\big\|^2_{L^2(\Omega_d)}\leq C\|\widetilde{\boldsymbol{v}}_d\|^2_{L^2(\Omega_d)}\leq C\|\widetilde{\boldsymbol{v}}_d\|^2_{H^1(\Omega_d)}, \tag{4.9}$$

$$\|Q_{d,0}\widetilde{\boldsymbol{v}}_d\|^2_{L^3(\Omega_d)}\leq C\big(\|Q_{d,0}\widetilde{\boldsymbol{v}}_d-\widetilde{\boldsymbol{v}}_d\|^2_{H^1(\Omega_d)}+\|\widetilde{\boldsymbol{v}}_d\|^2_{H^1(\Omega_d)}\big)\leq C\|\widetilde{\boldsymbol{v}}_d\|^2_{H^1(\Omega_d)}. \tag{4.10}$$

Thus combining (4.4)-(4.10), we can conclude that

$$
\begin{aligned}
|\!|\!| Q_h\widetilde{\boldsymbol{v}} |\!|\!|^2 &= 2\mu \sum_{T_s\in\mathcal{T}_{s,h}} \|\boldsymbol{D}_w(Q_{s,h}\widetilde{\boldsymbol{v}}_s)\|^2_{L^2(T_s)} + \frac{\mu\alpha}{\sqrt{\kappa}}\|Q_{s,b}\widetilde{\boldsymbol{v}}_s\cdot\boldsymbol{\tau}\|^2_{L^2(\Gamma_I)} \\
&\quad + \rho_s \sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\|Q_{s,0}\widetilde{\boldsymbol{v}}_s - Q_{s,b}\widetilde{\boldsymbol{v}}_s\|^2_{L^2(\partial T_s)} \\
&\quad + \rho_d \sum_{T_d\in\mathcal{T}_{d,h}} h_{T_d}^{-1}\|(Q_{d,0}\widetilde{\boldsymbol{v}}_d - Q_{d,b}\widetilde{\boldsymbol{v}}_d)\cdot\boldsymbol{n}_d\|^2_{L^2(\partial T_d)} \\
&\quad + \mu\|\mathbf{K}^{-\frac{1}{2}}Q_{d,0}\widetilde{\boldsymbol{v}}_d\|^2_{L^2(\Omega_d)} + \beta\rho\|Q_{d,0}\widetilde{\boldsymbol{v}}_d\|^2_{L^3(\Omega_d)} + \|\nabla_w\cdot(Q_h\widetilde{\boldsymbol{v}})\|^2_{L^2(\Omega)} \\
&\le C\big(\|\widetilde{\boldsymbol{v}}_s\|^2_{H^1(\Omega_s)} + \|\widetilde{\boldsymbol{v}}_d\|^2_{H^1(\Omega_d)}\big) \le C\big(\|\widetilde{\boldsymbol{v}}_s\|_{H^1(\Omega_s)} + \|\widetilde{\boldsymbol{v}}_d\|_{H^1(\Omega_d)}\big)^2.
\end{aligned}
$$

It follows from (A.2) and the definition of $\pi_{s,h}, \pi_{d,h}$ that

$$
\begin{aligned}
b_h(\boldsymbol{v}_h, p_h) &= (p_{s,h}, \nabla_w\cdot\boldsymbol{v}_{s,h})_{\Omega_s} + (p_{d,h}, \nabla_w\cdot\boldsymbol{v}_{d,h})_{\Omega_d} \\
&= \big(p_{s,h}, \nabla_w\cdot(Q_{s,h}\widetilde{\boldsymbol{v}}_s)\big)_{\Omega_s} + \big(p_{d,h}, \nabla_w\cdot(Q_{d,h}\widetilde{\boldsymbol{v}}_d)\big)_{\Omega_d} \\
&= \big(p_{s,h}, \pi_{s,h}(\nabla\cdot\widetilde{\boldsymbol{v}}_s)\big)_{\Omega_s} + \big(p_{d,h}, \pi_{d,h}(\nabla\cdot\widetilde{\boldsymbol{v}}_d)\big)_{\Omega_d} \\
&= (p_{s,h}, \nabla\cdot\widetilde{\boldsymbol{v}}_s)_{\Omega_s} + (p_{d,h}, \nabla\cdot\widetilde{\boldsymbol{v}}_d)_{\Omega_d}.
\end{aligned} \tag{4.11}
$$

Consequently, from (4.11), (4.3) and (4.2), we obtain

$$
\begin{aligned}
\frac{b_h(\boldsymbol{v}_h, p_h)}{\|p_h\|_h |\!|\!|\boldsymbol{v}_h|\!|\!|} &\ge \frac{(p_{s,h}, \nabla\cdot\widetilde{\boldsymbol{v}}_s)_{\Omega_s} + (p_{d,h}, \nabla\cdot\widetilde{\boldsymbol{v}}_d)_{\Omega_d}}{C\|p_h\|_h(\|\widetilde{\boldsymbol{v}}_s\|_{H^1(\Omega_s)} + \|\widetilde{\boldsymbol{v}}_d\|_{H^1(\Omega_d)})} \\
&\ge \frac{(p_{s,h}, p_{s,h})_{\Omega_s} + (p_{d,h}, p_{d,h})_{\Omega_d}}{C(\|p_{s,h}\|_{\Omega_s} + \|p_{d,h}\|_{\Omega_d})^2} \\
&= \frac{\|p_{s,h}\|^2_{\Omega_s} + \|p_{d,h}\|^2_{\Omega_d}}{C(\|p_{s,h}\|_{\Omega_s} + \|p_{d,h}\|_{\Omega_d})^2} \ge \frac{1}{2C}.
\end{aligned}
$$

The proof is complete. □

Finally, in order to illustrate the existence and uniqueness of the nonlinear system (3.1), we need to verify the following theorem.

**Lemma 4.2 (Minty-Browder Theorem, [12]).** *Let $V$ be a real separable reflexive Banach space. Denote a norm in $V$ as $\|\cdot\|_V$, and $\langle\cdot,\cdot\rangle_V$ is the duality between $(V)^*$ and $V$. Let $A: V \to V^*$ satisfy the following conditions:*

*(i) The operator $A$ is coercive:*

$$
\lim_{\boldsymbol{v}\in V, \|\boldsymbol{v}\|_V\to\infty} \frac{\langle A(\boldsymbol{v}), \boldsymbol{v}\rangle_V}{\|\boldsymbol{v}\|_V} = \infty.
$$

*(ii) $A$ is hemi-continuous: For $\boldsymbol{u}, \boldsymbol{v} \in V$ and $t \in R, t \to \langle A(\boldsymbol{u}+t\boldsymbol{v}), \boldsymbol{v}\rangle_V$ is a continuous mapping from $R$ to $R$.*

*(iii) $A$ is a strictly monotone operator:*

$$
\langle A(\boldsymbol{u}) - A(\boldsymbol{v}), \boldsymbol{u}-\boldsymbol{v}\rangle_V > 0, \quad \boldsymbol{u}, \boldsymbol{v}\in V, \quad \boldsymbol{u}\neq\boldsymbol{v}.
$$

*Then $A$ is a bijection, i.e. given any $F \in V^*$, there exists $\boldsymbol{u}$ such that*

$$\boldsymbol{u} \in V \quad \textit{and} \quad A(\boldsymbol{u}) = F.$$

**Lemma 4.3.** *The Weak Galerkin Algorithm* 3.1 *has a unique solution.*

*Proof.* Since $V_h$ is a real finite dimension Banach space, $V_h$ is separable and reflexive. It is easy to see that there exists a mapping $A : V_h \to (V_h)^*$, a functional $F \in (V_h)^*$ such that

$$\begin{aligned}\langle A(\boldsymbol{u}_h), \boldsymbol{v}_h\rangle_{V_h} &= a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) + s(\boldsymbol{u}_h, \boldsymbol{v}_h),\\ \langle F, \boldsymbol{v}_h\rangle_{V_h} &= (\boldsymbol{f}, \boldsymbol{v}_0), \quad \forall\, \boldsymbol{v}_h \in V_h.\end{aligned}$$

Here $(V_h)^*$ denotes the dual space of $V_h$.

Let

$$K(g) = \{\boldsymbol{u}_h \in V_h,\ b_h(\boldsymbol{u}_h, p_h) = (g, p_h),\ \forall\, p_h \in W_h\}.$$

We claim that if $\boldsymbol{v}_h \in K(0)$, then $\nabla_w \cdot \boldsymbol{v}_h = 0$. In fact, by the definition of $V_h$ and Definition 3.1, we know that for any $\boldsymbol{v}_h \in V_h$,

$$\begin{aligned}(\nabla_w \cdot \boldsymbol{v}_{s,h}, 1)_{\Omega_s} + (\nabla_w \cdot \boldsymbol{v}_{d,h}, 1)_{\Omega_d} &= \langle \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s, 1\rangle_{\partial\Omega_s} + \langle \boldsymbol{v}_{d,b} \cdot \boldsymbol{n}_d, 1\rangle_{\partial\Omega_d}\\ &= \langle \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s + \boldsymbol{v}_{d,b} \cdot \boldsymbol{n}_d, 1\rangle_{\Gamma_I} = 0,\end{aligned}$$

which means $\nabla_w \cdot \boldsymbol{v}_h \in W_h$. Take $p_h = \nabla_w \cdot \boldsymbol{v}_h$, thus $b_h(\boldsymbol{v}_h, p_h) = (\nabla_w \cdot \boldsymbol{v}_h, \nabla_w \cdot \boldsymbol{v}_h) = 0$, the assertion holds. Therefore, we see that the WG scheme (3.1) can be written as in the following way, find $\boldsymbol{u}_h \in K(g)$ such that

$$\langle A(\boldsymbol{u}_h), \boldsymbol{v}_h\rangle_{V_h} = \langle F, \boldsymbol{v}_h\rangle_{V_h}, \quad \forall\, \boldsymbol{v}_h \in K(0). \tag{4.12}$$

There exists a mapping $B$, such that for $\boldsymbol{u}_h \in K(g)$, $B\boldsymbol{u}_h = g$. From [42, Lemma 3.5] and the inf-sup condition (4.1), there exists a function $\boldsymbol{u}_h^0 \in K(g)/\ker(B)$ such that $|||\boldsymbol{u}_h^0||| \le \|g\|/\varrho$. Consequently, it is clear that the scheme (4.12) is equivalent to the following scheme: Find $\hat{\boldsymbol{u}}_h = \boldsymbol{u}_h - \boldsymbol{u}_h^0 \in K(0)$ such that

$$\big\langle A\big(\hat{\boldsymbol{u}}_h + \boldsymbol{u}_h^0\big), \boldsymbol{v}_h\big\rangle_{V_h} = \langle F, \boldsymbol{v}_h\rangle_{V_h}, \quad \forall\, \boldsymbol{v}_h \in K(0). \tag{4.13}$$

The scheme (4.13) can be written as the operator equation $A(\hat{\boldsymbol{u}}_h + \boldsymbol{u}_h^0) = F$ for $\hat{\boldsymbol{u}}_h \in K(0)$.

According to Lemma 4.2, we need to show that for fixed $\boldsymbol{u}_h^0 \in V_h$, and $\forall\, \hat{\boldsymbol{u}}_h, \boldsymbol{v}_h \in K(0)$,

(i) The mapping $\hat{\boldsymbol{u}}_h \to A(\boldsymbol{u}_h^0 + \hat{\boldsymbol{u}}_h)$ is coercive

$$\lim_{|||\hat{\boldsymbol{u}}_h||| \to \infty} \frac{\big\langle A\big(\boldsymbol{u}_h^0 + \hat{\boldsymbol{u}}_h\big), \hat{\boldsymbol{u}}_h\big\rangle_{V_h}}{|||\hat{\boldsymbol{u}}_h|||} = \infty. \tag{4.14}$$

(ii) The mapping $A$ is hemi-continuous in $V_h$: The mapping

$$t \to \big\langle A\big(\boldsymbol{u}_h^0 + \hat{\boldsymbol{u}}_h + t\boldsymbol{v}_h\big), \boldsymbol{v}_h\big\rangle_{V_h} \tag{4.15}$$

is continuous from $R$ to $R$.

(iii) The mapping $\hat{\boldsymbol{u}}_h \to A(\boldsymbol{u}_h^0 + \hat{\boldsymbol{u}}_h)$ is monotone

$$\big\langle A\big(\hat{\boldsymbol{u}}_h + \boldsymbol{u}_h^0\big) - A\big(\boldsymbol{v}_h + \boldsymbol{u}_h^0\big), \hat{\boldsymbol{u}}_h - \boldsymbol{v}_h\big\rangle_{V_h} > 0, \quad \hat{\boldsymbol{u}}_h \ne \boldsymbol{v}_h. \tag{4.16}$$

Next, we divide our proof into three steps. First, we verify the coercivity of $A$. Let $\boldsymbol{u}_h^0+\hat{\boldsymbol{u}}_h=\overline{\boldsymbol{u}}_h$, according to the definition of $|\!|\!|\cdot|\!|\!|$ and Hölder inequality, it follows that

$$
\begin{aligned}
&\big\langle A\big(\boldsymbol{u}_h^0+\hat{\boldsymbol{u}}_h\big),\hat{\boldsymbol{u}}_h\big\rangle_{V_h}\\
&=a_h\big(\overline{\boldsymbol{u}}_h,\overline{\boldsymbol{u}}_h-\boldsymbol{u}_h^0\big)+s\big(\overline{\boldsymbol{u}}_h,\overline{\boldsymbol{u}}_h-\boldsymbol{u}_h^0\big)=a_h\big(\overline{\boldsymbol{u}}_h,\overline{\boldsymbol{u}}_h\big)+s\big(\overline{\boldsymbol{u}}_h,\overline{\boldsymbol{u}}_h\big)-a_h\big(\overline{\boldsymbol{u}}_h,\boldsymbol{u}_h^0\big)-s\big(\overline{\boldsymbol{u}}_h,\boldsymbol{u}_h^0\big)\\
&\geq 2\mu\sum_{T_s\in\mathcal{T}_{s,h}}\|\boldsymbol{D}_w(\overline{\boldsymbol{u}}_{s,h})\|^2_{L^2(T_s)}+\frac{\mu\alpha}{\sqrt{\kappa}}\|\overline{\boldsymbol{u}}_{s,b}\cdot\boldsymbol{\tau}\|^2_{L^2(\Gamma_I)}+\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\|\overline{\boldsymbol{u}}_{s,0}-\overline{\boldsymbol{u}}_{s,b}\|^2_{L^2(\partial T_s)}\\
&\quad+\sum_{T_d\in\mathcal{T}_{d,h}}h_{T_d}^{-1}\|(\overline{\boldsymbol{u}}_{d,0}-\overline{\boldsymbol{u}}_{d,b})\cdot\boldsymbol{n}_d\|^2_{L^2(\partial T_d)}+\mu\big\|\mathbf{K}^{-\frac12}\overline{\boldsymbol{u}}_{d,0}\big\|^2_{L^2(\Omega_d)}+\beta\rho\|\overline{\boldsymbol{u}}_{d,0}\|^3_{L^3(\Omega_d)}\\
&\quad-2\mu\Big(\sum_{T_s\in\mathcal{T}_{s,h}}\|\boldsymbol{D}_w(\overline{\boldsymbol{u}}_{s,h})\|^2_{L^2(T_s)}\Big)^{\frac12}\Big(\sum_{T_s\in\mathcal{T}_{s,h}}\big\|\boldsymbol{D}_w\big(\boldsymbol{u}_{s,h}^0\big)\big\|^2_{L^2(T_s)}\Big)^{\frac12}\\
&\quad-\frac{\mu\alpha}{\sqrt{\kappa}}\big(\|\overline{\boldsymbol{u}}_{s,b}\cdot\boldsymbol{\tau}\|^2_{L^2(\Gamma_I)}\big)^{\frac12}\Big(\big\|\boldsymbol{u}_{s,b}^0\cdot\boldsymbol{\tau}\big\|^2_{L^2(\Gamma_I)}\Big)^{\frac12}\\
&\quad-\Big(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\|\overline{\boldsymbol{u}}_{s,0}-\overline{\boldsymbol{u}}_{s,b}\|^2_{L^2(\partial T_s)}\Big)^{\frac12}\Big(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\big\|\boldsymbol{u}_{s,0}^0-\boldsymbol{u}_{s,b}^0\big\|^2_{L^2(\partial T_s)}\Big)^{\frac12}\\
&\quad-\Big(\sum_{T_d\in\mathcal{T}_{d,h}}h_{T_d}^{-1}\|(\overline{\boldsymbol{u}}_{d,0}-\overline{\boldsymbol{u}}_{d,b})\cdot\boldsymbol{n}_d\|^2_{L^2(\partial T_d)}\Big)^{\frac12}\Big(\sum_{T_d\in\mathcal{T}_{d,h}}h_{T_d}^{-1}\big\|\big(\boldsymbol{u}_{d,0}^0-\boldsymbol{u}_{d,b}^0\big)\cdot\boldsymbol{n}_d\big\|^2_{L^2(\partial T_d)}\Big)^{\frac12}\\
&\quad-\mu\big\|\boldsymbol{K}^{-\frac12}\overline{\boldsymbol{u}}_{d,0}\big\|_{L^2(\Omega_d)}\big\|\boldsymbol{K}^{-\frac12}\boldsymbol{u}_{d,0}^0\big\|_{L^2(\Omega_d)}-\beta\rho\|\overline{\boldsymbol{u}}_{d,0}\|^2_{L^3(\Omega_d)}\big\|\boldsymbol{u}_{d,0}^0\big\|_{L^3(\Omega_d)}\\
&\quad+\|\nabla_w\cdot\overline{\boldsymbol{u}}_h\|^2_{L^2(\Omega)}-\|\nabla_w\cdot\overline{\boldsymbol{u}}_h\|_{L^2(\Omega)}\big\|\nabla_w\cdot\boldsymbol{u}_h^0\big\|_{L^2(\Omega)}\\
&\geq\min\big\{1,\|\overline{\boldsymbol{u}}_{d,0}\|_{L^3(\Omega_d)}\big\}|\!|\!|\overline{\boldsymbol{u}}_h|\!|\!|^2-\max\Big\{1,\frac{1}{\beta\rho\|\boldsymbol{u}_{d,0}^0\|_{L^3(\Omega_d)}}\Big\}|\!|\!|\overline{\boldsymbol{u}}_h|\!|\!|\cdot|\!|\!|\boldsymbol{u}_h^0|\!|\!|,
\end{aligned}
$$

where we have used the fact that

$$
\|\nabla_w\cdot\overline{\boldsymbol{u}}_h\|_{L^2(\Omega)}\leq\big\|\nabla_w\cdot\boldsymbol{u}_h^0\big\|_{L^2(\Omega)}+\|\nabla_w\cdot\hat{\boldsymbol{u}}_h\|_{L^2(\Omega)}=\big\|\nabla_w\cdot\boldsymbol{u}_h^0\big\|_{L^2(\Omega)}.
$$

Since $\boldsymbol{u}_h^0$ is fixed, the factor

$$
\min\big\{1,\|\overline{\boldsymbol{u}}_{d,0}\|_{L^3(\Omega_d)}\big\}|\!|\!|\overline{\boldsymbol{u}}_h|\!|\!|-\max\Big\{1,\frac{1}{\beta\rho\|\boldsymbol{u}_{d,0}^0\|_{L^3(\Omega_d)}}\Big\}|\!|\!|\boldsymbol{u}_h^0|\!|\!|
$$

is nonnegative for $|\!|\!|\overline{\boldsymbol{u}}_h|\!|\!|$ large enough. In this case, when $|\!|\!|\hat{\boldsymbol{u}}_h|\!|\!|\to\infty$, (4.14) holds, which completes the proof of the coercivity.

Next, we prove the hemi-continuity of $A$. Let

$$
\boldsymbol{u}_h^0+\hat{\boldsymbol{u}}_h+t\boldsymbol{v}_h=\boldsymbol{w}_h,\quad \boldsymbol{u}_h^0+\hat{\boldsymbol{u}}_h+t_0\boldsymbol{v}_h=\boldsymbol{w}_h^0.
$$

By the definition of $|\!|\!|\cdot|\!|\!|$ and the Hölder inequality, we have

$$
\begin{aligned}
&a_s(\boldsymbol{w}_h,\boldsymbol{v}_h)+s(\boldsymbol{w}_h,\boldsymbol{v}_h)-a_s\big(\boldsymbol{w}_h^0,\boldsymbol{v}_h\big)-s\big(\boldsymbol{w}_h^0,\boldsymbol{v}_h\big)\\
&=a_s\big(\boldsymbol{w}_h-\boldsymbol{w}_h^0,\boldsymbol{v}_h\big)+s\big(\boldsymbol{w}_h-\boldsymbol{w}_h^0,\boldsymbol{v}_h\big)\\
&\leq 2\mu\Big(\sum_{T_s\in\mathcal{T}_{s,h}}\big\|\boldsymbol{D}_w\big(\boldsymbol{w}_{s,h}-\boldsymbol{w}_{s,h}^0\big)\big\|^2_{L^2(T_s)}\Big)^{\frac12}\Big(\sum_{T_s\in\mathcal{T}_{s,h}}\|\boldsymbol{D}_w(\boldsymbol{v}_{s,h})\|^2_{L^2(T_s)}\Big)^{\frac12}
\end{aligned}
$$

$$
\begin{aligned}
&+\Big(\rho_s \sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\big\|\big(\boldsymbol{w}_{s,0}-\boldsymbol{w}_{s,0}^0\big)-\big(\boldsymbol{w}_{s,b}-\boldsymbol{w}_{s,b}^0\big)\big\|_{L^2(\partial T_s)}^2\Big)^{\frac12}\Big(\rho_s \sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\|\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\|_{L^2(\partial T_s)}^2\Big)^{\frac12}\\
&+\Big(\rho_d \sum_{T_d\in\mathcal{T}_{d,h}} h_{T_d}^{-1}\big\|\big(\big(\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big)-\big(\boldsymbol{w}_{d,b}-\boldsymbol{w}_{d,b}^0\big)\big)\cdot\boldsymbol{n}_d\big\|_{L^2(\partial T_d)}^2\Big)^{\frac12}\\
&\times\Big(\rho_d \sum_{T_d\in\mathcal{T}_{d,h}} h_{T_d}^{-1}\|(\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b})\cdot\boldsymbol{n}_d\|_{L^2(\partial T_d)}^2\Big)^{\frac12}\\
&+\frac{\mu\alpha}{\sqrt{\kappa}}\Big(\big\|\big(\boldsymbol{w}_{s,b}-\boldsymbol{w}_{s,b}^0\big)\cdot\boldsymbol{\tau}\big\|_{L^2(\Gamma_I)}^2\Big)^{\frac12}\Big(\|\boldsymbol{v}_{s,b}\cdot\boldsymbol{\tau}\|_{L^2(\Gamma_I)}^2\Big)^{\frac12}\\
\le{}& C\big\|\!\big|\boldsymbol{w}_h-\boldsymbol{w}_h^0\big\|\!\big|_1 \|\!|\boldsymbol{v}_h\|\!|_1 \le C\big\|\!\big|\boldsymbol{w}_h-\boldsymbol{w}_h^0\big\|\!\big|\cdot\|\!|\boldsymbol{v}_h\|\!|. \qquad (4.17)
\end{aligned}
$$

From the proof of [27, Lemma 3.16] and Hölder inequality, we have

$$
\begin{aligned}
&\big|\beta\rho\big(|\boldsymbol{w}_{d,0}|\boldsymbol{w}_{d,0}-\big|\boldsymbol{w}_{d,0}^0\big|\boldsymbol{w}_{d,0}^0,\boldsymbol{v}_{d,0}\big)_{\Omega_d}\big|\\
\le{}& \beta\rho\int_{\Omega_d}\big|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big|\big(|\boldsymbol{w}_{d,0}|+\big|\boldsymbol{w}_{d,0}^0\big|\big)|\boldsymbol{v}_{d,0}|\mathrm{d}s\\
\le{}& \beta\rho\big\|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big\|_{L^{p_1}(\Omega_d)}\Big(\|\boldsymbol{w}_{d,0}\|_{L^{p_2}(\Omega_d)}+\big\|\boldsymbol{w}_{d,0}^0\big\|_{L^{p_2}(\Omega_d)}\Big)\|\boldsymbol{v}_{d,0}\|_{L^{p_3}(\Omega_d)}\\
\le{}& \beta\rho\big\|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big\|_{L^{p_1}(\Omega_d)}\Big(\big\|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big\|_{L^{p_2}(\Omega_d)}+2\big\|\boldsymbol{w}_{d,0}^0\big\|_{L^{p_2}(\Omega_d)}\Big)\|\boldsymbol{v}_{d,0}\|_{L^{p_3}(\Omega_d)}. \qquad (4.18)
\end{aligned}
$$

where $p_i > 1, i = 1,2,3$, satisfying $1/p_1 + 1/p_2 + 1/p_3 = 1$. Let $p_1 = 2, p_2 = 6, p_3 = 3$. We obtain

$$
\begin{aligned}
&a_d(\boldsymbol{w}_h,\boldsymbol{v}_h)-a_d\big(\boldsymbol{w}_h^0,\boldsymbol{v}_h\big)\\
={}& \mu\big(\boldsymbol{K}^{-1}\big(\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big),\boldsymbol{v}_{d,0}\big)_{\Omega_d}+\big(\beta\rho|\boldsymbol{w}_{d,0}|\boldsymbol{w}_{d,0}-\big|\boldsymbol{w}_{d,0}^0\big|\boldsymbol{w}_{d,0}^0,\boldsymbol{v}_{d,0}\big)_{\Omega_d}\\
\le{}& \mu\big\|\boldsymbol{K}^{-\frac12}\big(\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big)\big\|_{L^2(\Omega_d)}\big\|\boldsymbol{K}^{-\frac12}\boldsymbol{v}_{d,0}\big\|_{L^2(\Omega_d)}\\
&+\beta\rho\big\|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big\|_{L^2(\Omega_d)}\Big(\big\|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big\|_{L^6(\Omega_d)}+2\big\|\boldsymbol{w}_{d,0}^0\big\|_{L^6(\Omega_d)}\Big)\|\boldsymbol{v}_{d,0}\|_{L^3(\Omega_d)}\\
\le{}& C\big\|\!\big|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big\|\!\big|\cdot\|\!|\boldsymbol{v}_h\|\!|\Big(1+\big\|\boldsymbol{w}_{d,0}-\boldsymbol{w}_{d,0}^0\big\|_{L^6(\Omega_d)}+2\big\|\boldsymbol{w}_{d,0}^0\big\|_{L^6(\Omega_d)}\Big). \qquad (4.19)
\end{aligned}
$$

Hence, when $t \to t_0$, according to the continuity of the norm $\|\!|\cdot\|\!|$, (4.17) and (4.19), we have

$$
\begin{aligned}
&\lim_{t\to t_0}\Big(\big\langle A\big(\boldsymbol{u}_h^0+\hat{\boldsymbol{u}}_h+t\boldsymbol{v}_h\big),\boldsymbol{v}_h\big\rangle_{V_h}-\big\langle A\big(\boldsymbol{u}_h^0+\hat{\boldsymbol{u}}_h+t_0\boldsymbol{v}_h\big),\boldsymbol{v}_h\big\rangle_{V_h}\Big)\\
={}&\lim_{t\to t_0}\big(a_h(\boldsymbol{w}_h,\boldsymbol{v}_h)+s(\boldsymbol{w}_h,\boldsymbol{v}_h)-a_h\big(\boldsymbol{w}_h^0,\boldsymbol{v}_h\big)-s\big(\boldsymbol{w}_h^0,\boldsymbol{v}_h\big)\big)=0.
\end{aligned}
$$

Therefore, the mapping (4.15) is continuous for $t \in R$. Furthermore, $A$ is a hemi-continuous operator.

Finally, we prove the (4.16). Obviously, $h(\boldsymbol{x}) = |\boldsymbol{x}|\boldsymbol{x}$ is increasing in $\mathbb{R}^2$, then it follows that

$$
\begin{aligned}
&\big\langle A\big(\hat{\boldsymbol{u}}_h+\boldsymbol{u}_h^0\big)-A\big(\boldsymbol{v}_h+\boldsymbol{u}_h^0\big),\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h\big\rangle_{V_h}\\
={}& a_s(\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h,\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h)+s(\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h,\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h)\\
&+\mu\big(\boldsymbol{K}^{-1}(\hat{\boldsymbol{u}}_{d,0}-\boldsymbol{v}_{d,0}),\hat{\boldsymbol{u}}_{d,0}-\boldsymbol{v}_{d,0}\big)_{\Omega_d}\\
&+\big(\beta\rho(|\hat{\boldsymbol{u}}_{d,0}|\hat{\boldsymbol{u}}_{d,0}-|\boldsymbol{v}_{d,0}|\boldsymbol{v}_{d,0}),\hat{\boldsymbol{u}}_{d,0}-\boldsymbol{v}_{d,0}\big)_{\Omega_d}\ge 0.
\end{aligned}
$$

Let us assume that

$$\left\langle A\left(\hat{\boldsymbol{u}}_h+\boldsymbol{u}_h^0\right)-A\left(\boldsymbol{v}_h+\boldsymbol{u}_h^0\right),\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h\right\rangle_{V_h}=0.$$

It is easy to see that $|\!|\!|\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h|\!|\!|=0$ and $\hat{\boldsymbol{u}}_b-\boldsymbol{v}_b=\mathbf{0}$ on $\partial\Omega$. This implies that $\hat{\boldsymbol{u}}_h-\boldsymbol{v}_h=\mathbf{0}$, which is contradict to $\hat{\boldsymbol{u}}_h\neq\boldsymbol{v}_h$. Thus we conclude that $A$ is a strictly monotone operator. According to Lemma 4.2, it follows that the scheme (4.13) has a unique solution. We complete the proof. □

## 5. Error Analysis

In this section, we derive the error equation and the order of convergence for Weak Galerkin Algorithm 3.1.

### 5.1. Error equation

Let the exact solution $\boldsymbol{u}=(\boldsymbol{u}_s,\boldsymbol{u}_d), p=(p_s,p_d)$ for (2.1)-(2.9). $\boldsymbol{u}_h=\{\boldsymbol{u}_0,\boldsymbol{u}_b\}$ and $p_h$ be the weak Galerkin finite element solution arising from (3.1). Define the error functions by

$$\boldsymbol{e}_h=Q_h\boldsymbol{u}-\boldsymbol{u}_h=\{Q_0\boldsymbol{u}-\boldsymbol{u}_0,Q_b\boldsymbol{u}-\boldsymbol{u}_b\},\quad \varepsilon_h=\pi_hp-p_h. \tag{5.1}$$

**Lemma 5.1 ([26]).** *For any $\boldsymbol{w}_s\in H^1(\Omega_s),\rho_l\in H^1(\Omega_l)$ $(l=s,d)$ and $\boldsymbol{v}_h=(\boldsymbol{v}_{s,h},\boldsymbol{v}_{d,h})\in V_h$. It follows that*

$$\begin{aligned}
&\left(\nabla_w(Q_{s,h}\boldsymbol{w}_s),\nabla_w\boldsymbol{v}_{s,h}\right)_{\Omega_s}=(\nabla\boldsymbol{w}_s,\nabla\boldsymbol{v}_{s,0})_{\Omega_s}-\sum_{T_s\in\mathcal{T}_{s,h}}\langle\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b},(\mathbb{Q}_{s,h}\nabla\boldsymbol{w}_s)\cdot\boldsymbol{n}_s\rangle_{\partial T_s},\\
&(\nabla_w\cdot\boldsymbol{v}_{l,h},\pi_{l,h}\rho_l)_{\Omega_l}=(\nabla\cdot\boldsymbol{v}_{l,0},\rho_l)_{\Omega_l}-\sum_{T_l\in\mathcal{T}_{l,h}}\langle\boldsymbol{v}_{l,0}-\boldsymbol{v}_{l,b},(\pi_{l,h}\rho_l)\cdot\boldsymbol{n}_l\rangle_{\partial T_l}.
\end{aligned}$$

With the help of the above lemma, we now present the following error equations.

**Lemma 5.2.** *Let $\boldsymbol{e}_h$ and $\varepsilon_h$ be the error defined in (5.1). Then, for any $\boldsymbol{v}_h\in V_h$ and $q_h\in W_h$, we have*

$$\begin{aligned}
&a_s(\boldsymbol{e}_h,\boldsymbol{v}_h)+a_d(Q_h\boldsymbol{u},\boldsymbol{v}_h)-a_d(\boldsymbol{u}_h,\boldsymbol{v}_h)+s(\boldsymbol{e}_h,\boldsymbol{v}_h)-b_h(\boldsymbol{v}_h,\varepsilon_h)\\
=\;&\varphi_{\boldsymbol{u},p}(\boldsymbol{v_h})+s(Q_h\boldsymbol{u},\boldsymbol{v}_h)+\mathcal{K}_{\boldsymbol{u}}(\boldsymbol{v}_h)+l_I(\boldsymbol{v}_h), &(5.2)\\
&b_h(\boldsymbol{e}_h,q_h)=0, &(5.3)
\end{aligned}$$

*where*

$$\begin{aligned}
\varphi_{\boldsymbol{u},p}(\boldsymbol{v}_h)=\;&2\mu\sum_{T_s\in\mathcal{T}_{s,h}}\left\langle\big(\boldsymbol{D}(\boldsymbol{u}_s)-\mathbb{Q}_{s,h}\boldsymbol{D}(\boldsymbol{u}_s)\big)\cdot\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\right\rangle_{\partial T_s}\\
&+\sum_{T_s\in\mathcal{T}_{s,h}}\langle(p_s-\pi_{s,h}p_s)\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\rangle_{\partial T_s}\\
&+\sum_{T_d\in\mathcal{T}_{d,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d},\\
\mathcal{K}_u(\boldsymbol{v}_h)=\;&\mu\big(\boldsymbol{K}^{-1}(Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_d),\boldsymbol{v}_{d,0}\big)_{\Omega_d}+\beta\rho(|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d-|\boldsymbol{u}_d|\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{\Omega_d},\\
l_I(\boldsymbol{v}_h)=\;&\frac{\mu\alpha}{\sqrt{\kappa}}\langle(Q_{s,b}\boldsymbol{u}_s-\boldsymbol{u}_s)\cdot\boldsymbol{\tau},\boldsymbol{v}_{s,b}\cdot\boldsymbol{\tau}\rangle_{\Gamma_I}.
\end{aligned}$$

*Proof.* First, testing (2.1) by $\boldsymbol{v}_{s,0}$ of $\boldsymbol{v}_h = (\{\boldsymbol{v}_{s,0}, \boldsymbol{v}_{s,b}\}, \{\boldsymbol{v}_{d,0}, \boldsymbol{v}_{d,b}\}) \in V_h$, we arrive at

$$
\begin{aligned}
(\boldsymbol{f}_s, \boldsymbol{v}_{s,0}) &= -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big(\nabla \cdot \boldsymbol{D}(\boldsymbol{u}_s), \boldsymbol{v}_{s,0}\big)_{T_s} + \sum_{T_s \in \mathcal{T}_{s,h}} (\nabla p_s, \boldsymbol{v}_{s,0})_{T_s} \\
&= -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \langle \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} \rangle_{\partial T_s} + 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big(\boldsymbol{D}(\boldsymbol{u}_s), \nabla \boldsymbol{v}_{s,0}\big)_{T_s} \\
&\quad + \sum_{T_s \in \mathcal{T}_{s,h}} \langle p_s \boldsymbol{n}_s, \boldsymbol{v}_{s,0} \rangle_{\partial T_s} - \sum_{T_s \in \mathcal{T}_{s,h}} (p_s, \nabla \cdot \boldsymbol{v}_{s,0})_{T_s} \\
&= -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \langle \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} - 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \langle \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{\partial T_s} \\
&\quad + 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big(\boldsymbol{D}(\boldsymbol{u}_s), \nabla \boldsymbol{v}_{s,0}\big)_{T_s} + \sum_{T_s \in \mathcal{T}_{s,h}} \langle p_s \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} \\
&\quad + \sum_{T_s \in \mathcal{T}_{s,h}} \langle p_s \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{\partial T_s} - \sum_{T_s \in \mathcal{T}_{s,h}} (p_s, \nabla \cdot \boldsymbol{v}_{s,0})_{T_s}.
\end{aligned} \tag{5.4}
$$

According to Lemma 5.1, we obtain

$$
\begin{aligned}
& -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \langle \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} + 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big(\boldsymbol{D}(\boldsymbol{u}_s), \nabla \boldsymbol{v}_{s,0}\big)_{T_s} \\
={}& -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big\langle \big(\boldsymbol{D}(\boldsymbol{u}_s) - \mathbb{Q}_{s,h} \boldsymbol{D}(\boldsymbol{u}_s)\big) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \big\rangle_{\partial T_s} \\
& -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big\langle \big(\mathbb{Q}_{s,h} \boldsymbol{D}(\boldsymbol{u}_s)\big) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \big\rangle_{\partial T_s} + 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big(\boldsymbol{D}(\boldsymbol{u}_s), \nabla \boldsymbol{v}_{s,0}\big)_{T_s} \\
={}& -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big\langle \big(\boldsymbol{D}(\boldsymbol{u}_s) - \mathbb{Q}_{s,h} \boldsymbol{D}(\boldsymbol{u}_s)\big) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \big\rangle_{\partial T_s} \\
& + 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big(\boldsymbol{D}_w(Q_{s,h} \boldsymbol{u}_s), \boldsymbol{D}_w(\boldsymbol{v}_{s,h})\big)_{T_s},
\end{aligned} \tag{5.5}
$$

$$
\begin{aligned}
& \sum_{T_s \in \mathcal{T}_{s,h}} \langle p_s \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} - \sum_{T_s \in \mathcal{T}_{s,h}} (p_s, \nabla \cdot \boldsymbol{v}_{s,0})_{T_s} \\
={}& \sum_{T_s \in \mathcal{T}_{s,h}} \langle (p_s - \pi_{s,h} p_s) \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} + \sum_{T_s \in \mathcal{T}_{s,h}} \langle (\pi_{s,h} p_s) \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} \\
& - \sum_{T_s \in \mathcal{T}_{s,h}} (p_s, \nabla \cdot \boldsymbol{v}_{s,0})_{T_s} \\
={}& \sum_{T_s \in \mathcal{T}_{s,h}} \langle (p_s - \pi_{s,h} p_s) \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} - \sum_{T_s \in \mathcal{T}_{s,h}} (\pi_{s,h} p_s, \nabla_w \cdot \boldsymbol{v}_{s,h})_{T_s}.
\end{aligned} \tag{5.6}
$$

Due to the Eqs. (2.8) and (2.9), we have

$$
\begin{aligned}
& -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \langle \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{\partial T_s} + \sum_{T_s \in \mathcal{T}_{s,h}} \langle p_s \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{\partial T_s} \\
={}& -2\mu \sum_{E_s \in \mathcal{E}_{I,h}} \langle \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{E_s} + \sum_{E_s \in \mathcal{E}_{I,h}} \langle p_s \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{E_s} \\
={}& -2\mu \sum_{E_s \in \mathcal{E}_{I,h}} \langle \boldsymbol{n}_s \cdot \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s \rangle_{E_s} - 2\mu \sum_{E_s \in \mathcal{E}_{I,h}} \langle \boldsymbol{n}_s \cdot \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s}
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{E_s \in \mathcal{E}_{I,h}} \langle p_s, \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s \rangle_{E_s} + \sum_{E_s \in \mathcal{E}_{I,h}} \langle p_s \boldsymbol{n}_s \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s} \\
=& \sum_{E_s \in \mathcal{E}_{I,h}} \langle p_d, \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s \rangle_{E_s} + \frac{\mu\alpha}{\sqrt{\kappa}} \sum_{E_s \in \mathcal{E}_{I,h}} \langle \boldsymbol{u}_s \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s} \\
=& \sum_{E_s \in \mathcal{E}_{I,h}} \langle p_d, \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s \rangle_{E_s} + \frac{\mu\alpha}{\sqrt{\kappa}} \sum_{E_s \in \mathcal{E}_{I,h}} \langle (\boldsymbol{u}_s - Q_{s,b}\boldsymbol{u}_s) \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s} \\
&+ \frac{\mu\alpha}{\sqrt{\kappa}} \sum_{E_s \in \mathcal{E}_{I,h}} \langle (Q_{s,b}\boldsymbol{u}_s) \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s},
\end{aligned}
\tag{5.7}
$$

where we have used the fact that

$$
\sum_{E_s \in \mathcal{E}_{s,h}^0 \cup \mathcal{E}_{s,h}^b} \langle \boldsymbol{D}(\boldsymbol{u}_s) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{E_s} = 0, \qquad \sum_{E_s \in \mathcal{E}_{s,h}^0 \cup \mathcal{E}_{s,h}^b} \langle p_s \boldsymbol{n}_s, \boldsymbol{v}_{s,b} \rangle_{E_s} = 0.
$$

Substituting (5.5)-(5.7) into (5.4), we get

$$
\begin{aligned}
(\boldsymbol{f}_s, \boldsymbol{v}_{s,0}) =& -2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big\langle \big(\boldsymbol{D}(\boldsymbol{u}_s) - \mathbb{Q}_{s,h}\boldsymbol{D}(\boldsymbol{u}_s)\big) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \big\rangle_{\partial T_s} \\
&+ 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big(\boldsymbol{D}_w(Q_{s,h}\boldsymbol{u}_s), \boldsymbol{D}_w(\boldsymbol{v}_{s,h})\big)_{T_s} \\
&+ \sum_{T_s \in \mathcal{T}_{s,h}} \langle (p_s - \pi_{s,h} p_s)\boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} \\
&- \sum_{T_s \in \mathcal{T}_{s,h}} (\pi_{s,h} p_s, \nabla_w \cdot \boldsymbol{v}_{s,h})_{T_s} + \sum_{E_s \in \mathcal{E}_{I,h}} \langle p_d, \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s \rangle_{E_s} \\
&+ \frac{\mu\alpha}{\sqrt{\kappa}} \sum_{E_s \in \mathcal{E}_{I,h}} \langle (\boldsymbol{u}_s - Q_{s,b}\boldsymbol{u}_s) \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s} \\
&+ \frac{\mu\alpha}{\sqrt{\kappa}} \sum_{E_s \in \mathcal{E}_{I,h}} \langle (Q_{s,b}\boldsymbol{u}_s) \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s} \\
=& \, a_s(Q_{s,h}\boldsymbol{u}_s, \boldsymbol{v}_{s,h}) - b_s(\boldsymbol{v}_{s,h}, \pi_h p_s) \\
&- 2\mu \sum_{T_s \in \mathcal{T}_{s,h}} \big\langle \big(\boldsymbol{D}(\boldsymbol{u}_s) - \mathbb{Q}_{s,h}\boldsymbol{D}(\boldsymbol{u}_s)\big) \cdot \boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \big\rangle_{\partial T_s} \\
&+ \sum_{T_s \in \mathcal{T}_{s,h}} \langle (p_s - \pi_{s,h} p_s)\boldsymbol{n}_s, \boldsymbol{v}_{s,0} - \boldsymbol{v}_{s,b} \rangle_{\partial T_s} + \sum_{E_s \in \mathcal{E}_{I,h}} \langle p_d, \boldsymbol{v}_{s,b} \cdot \boldsymbol{n}_s \rangle_{E_s} \\
&+ \frac{\mu\alpha}{\sqrt{\kappa}} \sum_{E_s \in \mathcal{E}_{I,h}} \langle (\boldsymbol{u}_s - Q_{s,b}\boldsymbol{u}_s) \cdot \boldsymbol{\tau}, \boldsymbol{v}_{s,b} \cdot \boldsymbol{\tau} \rangle_{E_s}.
\end{aligned}
\tag{5.8}
$$

Next, testing the nonlinear boundary condition (2.4) by $\boldsymbol{v}_{d,0}$, we have

$$
\begin{aligned}
(\boldsymbol{f}_d, \boldsymbol{v}_{d,0}) =& \, \mu \sum_{T_d \in \mathcal{T}_{d,h}} \big(\boldsymbol{K}^{-1}\boldsymbol{u}_d, \boldsymbol{v}_{d,0}\big)_{T_d} + \beta\rho \sum_{T_d \in \mathcal{T}_{d,h}} (|\boldsymbol{u}_d|\boldsymbol{u}_d, \boldsymbol{v}_{d,0})_{T_d} + \sum_{T_d \in \mathcal{T}_{d,h}} (\nabla p_d, \boldsymbol{v}_{d,0})_{T_d} \\
=& \, \mu \sum_{T_d \in \mathcal{T}_{d,h}} \big(\boldsymbol{K}^{-1}(Q_{d,0}\boldsymbol{u}_d), \boldsymbol{v}_{d,0}\big)_{T_d} + \mu \sum_{T_d \in \mathcal{T}_{d,h}} \big(\boldsymbol{K}^{-1}(\boldsymbol{u}_d - Q_{d,0}\boldsymbol{u}_d), \boldsymbol{v}_{d,0}\big)_{T_d} \\
&+ \beta\rho \sum_{T_d \in \mathcal{T}_{d,h}} (|Q_{d,0}\boldsymbol{u}_d| Q_{d,0}\boldsymbol{u}_d, \boldsymbol{v}_{d,0})_{T_d} + \beta\rho \sum_{T_d \in \mathcal{T}_{d,h}} (|\boldsymbol{u}_d|\boldsymbol{u}_d - |Q_{d,0}\boldsymbol{u}_d| Q_{d,0}\boldsymbol{u}_d, \boldsymbol{v}_{d,0})_{T_d} \\
&+ \sum_{T_d \in \mathcal{T}_{d,h}} \langle p_d \boldsymbol{n}_d, \boldsymbol{v}_{d,0} - \boldsymbol{v}_{d,b} \rangle_{\partial T_d} + \sum_{T_d \in \mathcal{T}_{d,h}} \langle p_d \boldsymbol{n}_d, \boldsymbol{v}_{d,b} \rangle_{\partial T_d} - \sum_{T_d \in \mathcal{T}_{d,h}} (p_d, \nabla \cdot \boldsymbol{v}_{d,0})_{T_d}.
\end{aligned}
\tag{5.9}
$$

From Lemma 5.1, we obtain

$$
\begin{aligned}
&\sum_{T_d\in\mathcal{T}_{d,h}}\langle p_d\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}-\sum_{T_d\in\mathcal{T}_{d,h}}(p_d,\nabla\cdot\boldsymbol{v}_{d,0})_{T_d}\\
=&\sum_{T_d\in\mathcal{T}_{d,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}+\sum_{T_d\in\mathcal{T}_{d,h}}\langle(\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}\\
&-\sum_{T_d\in\mathcal{T}_{d,h}}(p_d,\nabla\cdot\boldsymbol{v}_{d,0})_{T_d}\\
=&\sum_{T_d\in\mathcal{T}_{s,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}-\sum_{T_d\in\mathcal{T}_{d,h}}(\pi_{d,h}p_d,\nabla_w\cdot\boldsymbol{v}_{d,h})_{T_d}. \qquad (5.10)
\end{aligned}
$$

According to the fact

$$
\sum_{E_d\in\mathcal{E}_{d,h}^0\cup\mathcal{E}_{d,h}^b}\langle p_d\boldsymbol{n}_d,\boldsymbol{v}_{d,b}\rangle_{E_d}=0,
$$

it is clear that

$$
\sum_{T_d\in\mathcal{T}_{d,h}}\langle p_d\boldsymbol{n}_d,\boldsymbol{v}_{d,b}\rangle_{\partial T_d}=\sum_{E_s\in\mathcal{E}_{I,h}}\langle p_d,\boldsymbol{v}_{d,b}\cdot\boldsymbol{n}_d\rangle_{E_s}. \qquad (5.11)
$$

Substituting (5.10) and (5.11) into (5.9), we have

$$
\begin{aligned}
(\boldsymbol{f}_d,\boldsymbol{v}_{d,0})=&\,\mu\sum_{T_d\in\mathcal{T}_{d,h}}\big(\boldsymbol{K}^{-1}(Q_{d,0}\boldsymbol{u}_d),\boldsymbol{v}_{d,0}\big)_{T_d}+\mu\sum_{T_d\in\mathcal{T}_{d,h}}\big(\boldsymbol{K}^{-1}(\boldsymbol{u}_d-Q_{d,0}\boldsymbol{u}_d),\boldsymbol{v}_{d,0}\big)_{T_d}\\
&+\beta\rho\sum_{T_d\in\mathcal{T}_{d,h}}(|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{T_d}+\beta\rho\sum_{T_d\in\mathcal{T}_{d,h}}(|\boldsymbol{u}_d|\boldsymbol{u}_d-|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{T_d}\\
&+\sum_{T_d\in\mathcal{T}_{s,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}-\sum_{T_d\in\mathcal{T}_{d,h}}(\pi_{d,h}p_d,\nabla_w\cdot\boldsymbol{v}_{d,h})_{T_d}\\
&+\sum_{E_s\in\mathcal{E}_{I,h}}\langle p_d,\boldsymbol{v}_{d,b}\cdot\boldsymbol{n}_d\rangle_{E_s}\\
=&\,a_d(Q_{h,d}\boldsymbol{u}_d,\boldsymbol{v}_{d,h})-b_d(\boldsymbol{v}_{d,h}\pi_{d,h}p_d)+\sum_{E_s\in\mathcal{E}_{I,h}}\langle p_d,\boldsymbol{v}_{d,b}\cdot\boldsymbol{n}_d\rangle_{E_s}\\
&+\mu\sum_{T_d\in\mathcal{T}_{d,h}}\big(\boldsymbol{K}^{-1}(\boldsymbol{u}_d-Q_{d,0}\boldsymbol{u}_d),\boldsymbol{v}_{d,0}\big)_{T_d}+\beta\rho\sum_{T_d\in\mathcal{T}_{d,h}}(|\boldsymbol{u}_d|\boldsymbol{u}_d-|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{T_d}\\
&+\sum_{T_d\in\mathcal{T}_{s,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}. \qquad (5.12)
\end{aligned}
$$

Combining (5.4), (5.12), and (2.7), we obtain

$$
\begin{aligned}
(\boldsymbol{f},\boldsymbol{v}_0)=&\,a_h(Q_h\boldsymbol{u},\boldsymbol{v}_h)-b_h(\boldsymbol{v}_h,\pi_hp)-2\mu\sum_{T_s\in\mathcal{T}_{s,h}}\big\langle\big(\boldsymbol{D}(\boldsymbol{u}_s)-\mathbb{Q}_{s,h}\boldsymbol{D}(\boldsymbol{u}_s)\big)\cdot\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\big\rangle_{\partial T_s}\\
&+\sum_{T_s\in\mathcal{T}_{s,h}}\langle(p_s-\pi_{s,h}p_s)\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\rangle_{\partial T_s}+\frac{\mu\alpha}{\sqrt{\kappa}}\sum_{E_s\in\mathcal{E}_{I,h}}\langle(\boldsymbol{u}_s-Q_{s,b}\boldsymbol{u}_s)\cdot\boldsymbol{\tau},\boldsymbol{v}_{s,b}\cdot\boldsymbol{\tau}\rangle_{E_s}\\
&+\mu\sum_{T_d\in\mathcal{T}_{d,h}}\big(\boldsymbol{K}^{-1}(\boldsymbol{u}_d-Q_{d,0}\boldsymbol{u}_d),\boldsymbol{v}_{d,0}\big)_{T_d}+\beta\rho\sum_{T_d\in\mathcal{T}_{d,h}}(|\boldsymbol{u}_d|\boldsymbol{u}_d-|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{T_d}\\
&+\sum_{T_d\in\mathcal{T}_{s,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}\\
=&\,a_h(Q_h\boldsymbol{u},\boldsymbol{v}_h)-b_h(\boldsymbol{v}_h,\pi_hp)-\varphi_{\boldsymbol{u},p}(\boldsymbol{v}_{\boldsymbol{h}})-\mathcal{K}_{\boldsymbol{u}}(\boldsymbol{v}_h)-l_I(\boldsymbol{v}_h). \qquad (5.13)
\end{aligned}
$$

Subtracting (5.13) from (3.1) and adding $s(Q_h\boldsymbol{u},\boldsymbol{v}_h)$ to both sides, we yield the following error equation:

$$\begin{aligned}&a_s(\boldsymbol{e}_h,\boldsymbol{v}_h)+a_d(Q_h\boldsymbol{u},\boldsymbol{v}_h)-a_d(\boldsymbol{u}_h,\boldsymbol{v}_h)+s(\boldsymbol{e}_h,\boldsymbol{v}_h)-b_h(\boldsymbol{v}_h,\varepsilon_h)\\ =\ &\varphi_{\boldsymbol{u},p}(\boldsymbol{v_h})+s(Q_h\boldsymbol{u},\boldsymbol{v}_h)+\mathcal{K}_{\boldsymbol{u}}(\boldsymbol{v}_h)+l_I(\boldsymbol{v}_h).\end{aligned}$$

Finally, testing (2.2) by $q_{s,h}$ and testing (2.5) by $q_{d,h}$, $q_h=(q_{s,h},q_{d,h})\in W_h$. Then, it follows from Lemma A.1 that

$$(g,q_h)=(\nabla\cdot\boldsymbol{u},q_h)=\big(\pi_h(\nabla\cdot\boldsymbol{u}),q_h\big)=(\nabla_w\cdot Q_h\boldsymbol{u},q_h). \tag{5.14}$$

From the WG scheme (3.1) we know

$$(\nabla_w\cdot\boldsymbol{u}_h,q_h)=(g,q_h). \tag{5.15}$$

Combining (5.14) and (5.15), it leads to

$$b_h(\boldsymbol{e}_h,q_h)=\big(\nabla_w\cdot(Q_h\boldsymbol{u}-\boldsymbol{u}_h),q_h\big)=0.$$

This completes the proof of the lemma. □

### 5.2. Error estimate

Next, we derive the order of convergence for the WG approximation in $|\!|\!|\cdot|\!|\!|_1$ for $\boldsymbol{e}_h$ and in $\|\cdot\|_h$ for $\varepsilon_h$. It has been proved in [19, Lemma 1] that

$$(|a|^\gamma a-|b|^\gamma b)(a-b)\ge|a-b|^{\gamma+2},\quad\forall\,a,b\in\mathbb{R}^2,\quad\gamma>0.$$

It is straightforward to show that

$$(|\boldsymbol{a}|\boldsymbol{a}-|\boldsymbol{b}|\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b})\ge|\boldsymbol{a}-\boldsymbol{b}|^3,\quad\forall\,\boldsymbol{a},\boldsymbol{b}\in\big[L^3(\Omega_d)\big]^2. \tag{5.16}$$

Let $\boldsymbol{v}_h=\boldsymbol{e}_h$, therefore $b_h(\boldsymbol{e}_h,\varepsilon_h)=0$. According to (5.16), it follows that

$$\begin{aligned}&\beta\rho\sum_{T_d\in\mathcal{T}_{d,h}}(|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d-|\boldsymbol{u}_d|\boldsymbol{u}_d,\boldsymbol{e}_{d,0})_{T_d}\\ \ge\ &\beta\rho\sum_{T_d\in\mathcal{T}_{d,h}}\|Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_{d,h}\|^3_{L^3(T_d)}\\ =\ &\beta\rho\sum_{T_d\in\mathcal{T}_{d,h}}(|\boldsymbol{e}_{d,0}|\boldsymbol{e}_{d,0},\boldsymbol{e}_{d,0})_{T_d}.\end{aligned}$$

Thus we conclude that

$$\begin{aligned}&a_s(\boldsymbol{e}_h,\boldsymbol{e}_h)+a_d(Q_h\boldsymbol{u},\boldsymbol{e}_h)-a_d(\boldsymbol{u}_h,\boldsymbol{e}_h)+s(\boldsymbol{e}_h,\boldsymbol{e}_h)-b_h(\boldsymbol{e}_h,\varepsilon_h)\\ \ge\ &a_h(\boldsymbol{e}_h,\boldsymbol{e}_h)+s(\boldsymbol{e}_h,\boldsymbol{e}_h)=|\!|\!|\boldsymbol{e}_h|\!|\!|_1^2+\|\boldsymbol{e}_{d,0}\|^3_{L^3(\Omega_d)}.\end{aligned} \tag{5.17}$$

**Lemma 5.3.** *Assume that $\mathcal{T}_h$ is shape-regular. For $\boldsymbol{u}=(\boldsymbol{u}_s,\boldsymbol{u}_d)\in[H^{k+1}(\Omega_s)]^2\times[H^{k+1}(\Omega_d)]^2$, $p=(p_s,p_d)\in H^k(\Omega_s)\times H^k(\Omega_d)$ and $\boldsymbol{v}_h=(\{\boldsymbol{v}_{s,0},\boldsymbol{v}_{s,b}\},\{\boldsymbol{v}_{d,0},\boldsymbol{v}_{d,b}\})\in V_h$, it follows that*

$$|\varphi_{\boldsymbol{u},p}(\boldsymbol{v}_h)|\le\ Ch^k\left(\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}+\|p_s\|_{H^k(\Omega_s)}+\|p_d\|_{H^k(\Omega_d)}\right)|\!|\!|\boldsymbol{v}_h|\!|\!|_1, \tag{5.18}$$

$$|s(Q_h\boldsymbol{u},\boldsymbol{v}_h)| \le\ Ch^k\left(\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}+\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}\right)|\!|\!|\boldsymbol{v}_h|\!|\!|_1, \tag{5.19}$$

$$|\mathcal{K}_{\boldsymbol{u}}(\boldsymbol{v}_h)| \le\ Ch^{k+1}\big(\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}+\|\boldsymbol{u}_d\|^2_{H^{k+1}(\Omega_d)}\big)|\!|\!|\boldsymbol{v}_h|\!|\!|, \tag{5.20}$$

$$|l_I(\boldsymbol{v}_h)| \le\ Ch^{k+1}\|\boldsymbol{u}_s\|_{H^{k+1}(\Gamma_I)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1, \tag{5.21}$$

*where*

$$\begin{aligned}\varphi_{\boldsymbol{u},p}(\boldsymbol{v}_h) = 2\mu &\sum_{T_s\in\mathcal{T}_{s,h}}\big\langle\big(\boldsymbol{D}(\boldsymbol{u}_s)-\mathbb{Q}_{s,h}\boldsymbol{D}(\boldsymbol{u}_s)\big)\cdot\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\big\rangle_{\partial T_s}\\ &+\sum_{T_s\in\mathcal{T}_{s,h}}\langle(p_s-\pi_{s,h}p_s)\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\rangle_{\partial T_s}+\sum_{T_d\in\mathcal{T}_{d,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d},\end{aligned}$$

$$\mathcal{K}_u(\boldsymbol{v}_h) = \mu\big(\boldsymbol{K}^{-1}(Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_d),\boldsymbol{v}_{d,0}\big)_{\Omega_d}+\beta\rho(|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d-|\boldsymbol{u}_d|\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{\Omega_d},$$

$$l_I(\boldsymbol{v}_h) = \frac{\mu\alpha}{\sqrt{\kappa}}\langle(Q_{s,b}\boldsymbol{u}_s-\boldsymbol{u}_s)\cdot\boldsymbol{\tau},\boldsymbol{v}_{s,b}\cdot\boldsymbol{\tau}\rangle_{\Gamma_I}.$$

*Proof.* First, we estimate $\varphi_{\boldsymbol{u},p}(\boldsymbol{v}_h)$. According to Lemmas A.2 and A.4, it follows that

$$\begin{aligned}&\left|2\mu\sum_{T_s\in\mathcal{T}_{s,h}}\big\langle\big(\boldsymbol{D}(\boldsymbol{u}_s)-\mathbb{Q}_{s,h}\boldsymbol{D}(\boldsymbol{u}_s)\big)\cdot\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\big\rangle_{\partial T_s}\right|\\ \le\ & C\left(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}\|\boldsymbol{D}(\boldsymbol{u}_s)-\mathbb{Q}_{s,h}\boldsymbol{D}(\boldsymbol{u}_s)\|^2_{L^2(\partial T_s)}\right)^{\frac12}\left(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\|\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\|^2_{L^2(\partial T_s)}\right)^{\frac12}\\ \le\ & C\left(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}\big(h_{T_s}^{-1}\|\nabla\boldsymbol{u}_s-\mathbb{Q}_{s,h}\nabla\boldsymbol{u}_s\|^2_{L^2(T_s)}+h_{T_s}\|\nabla\boldsymbol{u}_s-\mathbb{Q}_{s,h}\nabla\boldsymbol{u}_s\|^2_{H^1(T_s)}\big)\right)^{\frac12}\\ &\times\left(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\|\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\|^2_{L^2(\partial T_s)}\right)^{\frac12}\\ \le\ & Ch^k\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1,\end{aligned}$$

and

$$\begin{aligned}&\left|\sum_{T_s\in\mathcal{T}_{s,h}}\langle(p_s-\pi_{s,h}p_s)\boldsymbol{n}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\rangle_{\partial T_s}\right|\\ \le\ & C\left(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}\|p_s-\pi_{s,h}p_s\|^2_{L^2(\partial T_s)}\right)^{\frac12}\left(\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\|\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\|^2_{L^2(\partial T_s)}\right)^{\frac12}\\ \le\ & Ch^k\|p_s\|_{H^k(\Omega_s)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1.\end{aligned}$$

Similarly,

$$\left|\sum_{T_d\in\mathcal{T}_{d,h}}\langle(p_d-\pi_{d,h}p_d)\boldsymbol{n}_d,\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b}\rangle_{\partial T_d}\right|\le Ch^k\|p_d\|_{H^k(\Omega_d)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1.$$

Thus we arrive at the conclusion that

$$|\varphi_{\boldsymbol{u},p}(\boldsymbol{v}_h)|\le Ch^k\left(\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}+\|p_s\|_{H^k(\Omega_s)}+\|p_d\|_{H^k(\Omega_d)}\right)|\!|\!|\boldsymbol{v}_h|\!|\!|_1.$$

Next, we estimate $s(Q_h\boldsymbol{u},\boldsymbol{v}_h)$. According to Lemmas A.2 and A.4, it is easy to see that

$$\left|\rho_s\sum_{T_s\in\mathcal{T}_{s,h}}h_{T_s}^{-1}\langle Q_{s,0}\boldsymbol{u}_s-Q_{s,b}\boldsymbol{u}_s,\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\rangle_{\partial T_s}\right|$$

$$
\begin{aligned}
&\leq C\left(\sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\|Q_{s,0}\boldsymbol{u}_s-Q_{s,b}\boldsymbol{u}_s\|_{L^2(\partial T_s)}^2\right)^{\frac{1}{2}}\left(\rho_s\sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\|\boldsymbol{v}_{s,0}-\boldsymbol{v}_{s,b}\|_{L^2(\partial T_s)}^2\right)^{\frac{1}{2}}\\
&\leq C\left(\sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\left(\|Q_{s,0}\boldsymbol{u}_s-\boldsymbol{u}_s\|_{L^2(\partial T_s)}+\|Q_{s,b}\boldsymbol{u}_s-\boldsymbol{u}_s\|_{L^2(\partial T_s)}\right)^2\right)^{\frac{1}{2}}|\!|\!|\boldsymbol{v}_h|\!|\!|_1\\
&\leq 2C\left(\sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\|Q_{s,0}\boldsymbol{u}_s-\boldsymbol{u}_s\|_{L^2(\partial T_s)}^2\right)^{\frac{1}{2}}|\!|\!|\boldsymbol{v}_h|\!|\!|_1\\
&\leq 2C\left(\sum_{T_s\in\mathcal{T}_{s,h}} h_{T_s}^{-1}\left(h_{T_s}^{-1}\|Q_{s,0}\boldsymbol{u}_s-\boldsymbol{u}_s\|_{L^2(T_s)}^2+h_{T_s}\|\nabla(Q_{s,0}\boldsymbol{u}_s-\boldsymbol{u}_s)\|_{L^2(T_s)}^2\right)\right)^{\frac{1}{2}}|\!|\!|\boldsymbol{v}_h|\!|\!|_1\\
&\leq 2Ch^k\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1.
\end{aligned}
$$

Similarly,

$$
\left|\rho_d\sum_{T_d\in\mathcal{T}_{d,h}} h_{T_d}^{-1}\langle(Q_{d,0}\boldsymbol{u}_d-Q_{d,b}\boldsymbol{u}_d)\cdot\boldsymbol{n}_d,(\boldsymbol{v}_{d,0}-\boldsymbol{v}_{d,b})\cdot\boldsymbol{n}_d\rangle_{\partial T_d}\right|\leq Ch^k\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1.
$$

Consequently, we infer that

$$
|s(Q_h\boldsymbol{u},\boldsymbol{v}_h)|\leq Ch^k\left(\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}+\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}\right)|\!|\!|\boldsymbol{v}_h|\!|\!|_1.
$$

Another step in the proof is to estimate $\mathcal{K}_{\boldsymbol{u}}(\boldsymbol{v}_h)$. By Lemma A.4, it follows that

$$
\begin{aligned}
\left|\mu\big(\boldsymbol{K}^{-1}(Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_d),\boldsymbol{v}_{d,0}\big)_{\Omega_d}\right| &\leq C\|Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_d\|_{L^2(\Omega_d)}\left\|\boldsymbol{K}^{-\frac{1}{2}}\boldsymbol{v}_{d,0}\right\|_{L^2(\Omega_d)}\\
&= C\left(\sum_{T_d\in\mathcal{T}_{d,h}}\|Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_d\|_{L^2(T_d)}^2\right)^{\frac{1}{2}}|\!|\!|\boldsymbol{v}_h|\!|\!|_1\\
&\leq Ch^{k+1}\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1. \qquad (5.22)
\end{aligned}
$$

From the (4.18), if $p_1=2,p_2=6,p_3=3$, using Lemmas A.4 and A.5, we obtain

$$
\begin{aligned}
&\left|\beta\rho(|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d-|\boldsymbol{u}_d|\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{\Omega_d}\right|\\
&\leq\beta\rho\|Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_d\|_{L^2(\Omega_d)}\left(\|Q_{d,0}\boldsymbol{u}_d\|_{L^6(\Omega_d)}+\|\boldsymbol{u}_d\|_{L^6(\Omega_d)}\right)\|\boldsymbol{v}_{d,0}\|_{L^3(\Omega_d)}\\
&\leq Ch^{k+1}\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}\left(\|Q_{d,0}\boldsymbol{u}_d-\boldsymbol{u}_d\|_{H^1(\Omega_d)}+2\|\boldsymbol{u}_d\|_{H^1(\Omega_d)}\right)|\!|\!|\boldsymbol{v}_h|\!|\!|\\
&\leq Ch^{k+1}\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}^2|\!|\!|\boldsymbol{v}_h|\!|\!|.
\end{aligned}
$$

Thus, (5.20) holds, and the following equation also holds:

$$
\left|\beta\rho(|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d-|\boldsymbol{u}_d|\boldsymbol{u}_d,\boldsymbol{v}_{d,0})_{\Omega_d}\right|\leq Ch^{k+1}\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}^2\|\boldsymbol{v}_{d,0}\|_{L^3(\Omega_d)}. \qquad (5.23)
$$

Finally, we estimate (5.21). From Lemma A.4, it follows that

$$
\begin{aligned}
|l_I(\boldsymbol{v}_h)| &= \left|\frac{\mu\alpha}{\sqrt{\kappa}}\langle(Q_{s,b}\boldsymbol{u}_s-\boldsymbol{u}_s)\cdot\boldsymbol{\tau},\boldsymbol{v}_{s,b}\cdot\boldsymbol{\tau}\rangle_{\Gamma_I}\right|\\
&\leq C\|Q_{s,b}\boldsymbol{u}_s-\boldsymbol{u}_s\|_{L^2(\Gamma_I)}\|\boldsymbol{v}_{s,b}\|_{L^2(\Gamma_I)}\\
&\leq Ch^{k+1}\|\boldsymbol{u}_s\|_{H^{k+1}(\Gamma_I)}|\!|\!|\boldsymbol{v}_h|\!|\!|_1.
\end{aligned}
$$

This completes the proof of Lemma 5.3. □

Now we are ready to estimate the error $|\!|\!|\boldsymbol{e}_h|\!|\!|_1$, $\|e_{d,0}\|_{L^3(\Omega_d)}$ and $\|\varepsilon_h\|_h$.

**Theorem 5.1.** *Let $\boldsymbol{u}_h \in V_h$ and $p_h \in W_h$ be the solution of* (3.1)*. Assume that the exact solution $\boldsymbol{u} \in [H^{k+1}(\Omega_s)]^2 \times [H^{k+1}(\Omega_d)]^2$ and $p \in H^k(\Omega_s) \times H^k(\Omega_d)$. Then there exists a constant $C$ such that*

$$\begin{aligned}
\|\!|\boldsymbol{e}_h|\!\|_1^2 + \|\boldsymbol{e}_{d,0}\|_{L^3(\Omega_d)}^3 \leq Ch^{\min\{2k,\frac{3}{2}k+\frac{3}{2}\}}\big(&\|\boldsymbol{u_s}\|_{H^{k+1}(\Omega_s)}^2 + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)}^2 + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)}^3 \\
&+ \|p_s\|_{H^k(\Omega_s)}^2 + \|p_d\|_{H^k(\Omega_d)}^2\big), \\
\|\varepsilon_h\|_h \leq Ch^{\min\{k,\frac{3}{4}k+\frac{3}{4}\}}\big(&\|\boldsymbol{u_s}\|_{H^{k+1}(\Omega_s)} + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)} + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)}^2 \\
&+ \|p_s\|_{H^k(\Omega_s)} + \|p_d\|_{H^k(\Omega_d)}\big).
\end{aligned}$$

*Proof.* According to Lemma 5.3, (5.22), (5.23), and Lemma A.6, we have

$$\begin{aligned}
&|\varphi_{\boldsymbol{u},p}(\boldsymbol{e}_h)| \leq \epsilon C\|\!|\boldsymbol{e}_h|\!\|_1^2 + \frac{C}{4\epsilon}h^{2k}\big(\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}^2 + \|p_s\|_{H^k(\Omega_s)}^2 + \|p_d\|_{H^k(\Omega_d)}^2\big), \\
&|s(Q_h\boldsymbol{u}, \boldsymbol{e}_h)| \leq \epsilon C\|\!|\boldsymbol{e}_h|\!\|_1^2 + \frac{C}{4\epsilon}h^{2k}\big(\|\boldsymbol{u}_s\|_{H^{k+1}(\Omega_s)}^2 + \|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}^2\big), \\
&|\mathcal{K}_{\boldsymbol{u}}(\boldsymbol{e}_h)| \leq \epsilon C\big(\|\!|\boldsymbol{e}_h|\!\|_1^2 + \|\boldsymbol{e}_{d,0}\|_{L^3(\Omega_d)}^3\big) + \frac{C}{4\epsilon}h^{2k+2}\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}^2 + \frac{2C}{3\sqrt{3\epsilon}}h^{\frac{3}{2}k+\frac{3}{2}}\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)}^3, \\
&|l_I(\boldsymbol{e}_h)| \leq \epsilon C\|\!|\boldsymbol{e}_h|\!\|_1^2 + \frac{C}{4\epsilon}h^{2k+2}\|\boldsymbol{u}_s\|_{H^{k+1}(\Gamma_I)}^2.
\end{aligned}$$

Consequently, from Lemma 5.2 and (5.17),

$$\begin{aligned}
\|\!|\boldsymbol{e}_h|\!\|_1^2 + \|\boldsymbol{e}_{d,0}\|_{L^3(\Omega_d)}^3 \leq Ch^{\min\{2k,\frac{3}{2}k+\frac{3}{2}\}}\big(&\|\boldsymbol{u_s}\|_{H^{k+1}(\Omega_s)}^2 + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)}^2 + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)}^3 \\
&+ \|p_s\|_{H^k(\Omega_s)}^2 + \|p_d\|_{H^k(\Omega_d)}^2\big).
\end{aligned}$$

Next, to estimate $\|\varepsilon\|_h$, we have from Lemma 5.2 that

$$\begin{aligned}
|b_h(\boldsymbol{v}_h, \varepsilon_h)| = \big|&a_s(\boldsymbol{e}_h, \boldsymbol{v}_h) + a_d(Q_h\boldsymbol{u}, \boldsymbol{v}_h) - a_d(\boldsymbol{u}_h, \boldsymbol{v}_h) + s(\boldsymbol{e}_h, \boldsymbol{v}_h) \\
&- \big(\varphi_{\boldsymbol{u},p}(\boldsymbol{v_h}) + s(Q_h\boldsymbol{u}, \boldsymbol{v}_h) + \mathcal{K}_{\boldsymbol{u}}(\boldsymbol{v}_h) + l_I(\boldsymbol{v}_h)\big)\big|,
\end{aligned}$$

According to (4.18) and Lemma A.7, if $p_1 = 2, p_2 = 6, p_3 = 3$, then

$$\begin{aligned}
&|a_d(Q_h\boldsymbol{u}, \boldsymbol{v}_h) - a_d(\boldsymbol{u}_h, \boldsymbol{v}_h)| \\
&\leq \big|\mu\big(\boldsymbol{K}^{-1}(Q_{d,0}\boldsymbol{u}_d - \boldsymbol{u}_{d,0}), \boldsymbol{v}_{d,0}\big)_{\Omega_d}\big| + |\beta\rho(|Q_{d,0}\boldsymbol{u}_d|Q_{d,0}\boldsymbol{u}_d - |\boldsymbol{u}_{d,0}|\boldsymbol{u}_{d,0}, \boldsymbol{v}_{d,0})_{\Omega_d}| \\
&\leq C\|\!|\boldsymbol{e}_h|\!\|_1\|\!|\boldsymbol{v}_h|\!\| + C\|Q_{d,0}\boldsymbol{u}_d - \boldsymbol{u}_d\|_{L^2(\Omega_d)}\left(\|Q_{d,0}\boldsymbol{u}_d\|_{L^6(\Omega_d)} + \|\boldsymbol{u}_{d,0}\|_{L^6(\Omega_d)}\right)\|\boldsymbol{v}_{d,0}\|_{L^3(\Omega_d)} \\
&\leq C\|\!|\boldsymbol{e}_h|\!\|_1\|\!|\boldsymbol{v}_h|\!\| + C\|\!|\boldsymbol{e}_h|\!\|_1\left(2\|Q_{d,0}\boldsymbol{u}_d\|_{L^6(\Omega_d)} + \|\boldsymbol{e}_{d,0}\|_{L^6(\Omega_d)}\right)\|\!|\boldsymbol{v}_h|\!\| \\
&\leq C\|\!|\boldsymbol{e}_h|\!\|_1\|\!|\boldsymbol{v}_h|\!\| + C\|\!|\boldsymbol{e}_h|\!\|_1\left(2\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)} + h_{T_d}^{-\frac{2}{3}}\|\boldsymbol{e}_{d,0}\|_{L^2(\Omega_d)}\right)\|\!|\boldsymbol{v}_h|\!\| \\
&\leq C\|\!|\boldsymbol{e}_h|\!\|_1\|\!|\boldsymbol{v}_h|\!\| + C\|\!|\boldsymbol{e}_h|\!\|_1\left(2\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)} + h^{-\frac{2}{3}}\|\!|\boldsymbol{e}_h|\!\|_1\right)\|\!|\boldsymbol{v}_h|\!\|. \qquad (5.24)
\end{aligned}$$

Using the Eqs. (4.17), (5.24) and Lemma 5.3, we arrive at

$$\begin{aligned}
|b_h(\boldsymbol{v}_h, \varepsilon_h)| \leq C\Big(&1 + 2\|\boldsymbol{u}_d\|_{H^{k+1}(\Omega_d)} + h^{-\frac{2}{3}}\|\!|\boldsymbol{e}_h|\!\|_1\Big)\|\!|\boldsymbol{e}_h|\!\|_1\|\!|\boldsymbol{v}_h|\!\| \\
&+ Ch^k\big(\|\boldsymbol{u_s}\|_{H^{k+1}(\Omega_s)} + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)} + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)}^2 \\
&+ \|p_s\|_{H^k(\Omega_s)} + \|p_d\|_{H^k(\Omega_d)}\big)\|\!|\boldsymbol{v}_h|\!\|.
\end{aligned}$$

Combining the above estimate with the inf-sup condition (4.1) gives

$$\|\varepsilon_h\|_h \leq \frac{b_h(\boldsymbol{v}_h, p_h)}{\varrho \| \boldsymbol{v}_h \|} \leq Ch^{\min\{k, \frac{3}{4}k+\frac{3}{4}\}} \big( \|\boldsymbol{u_s}\|_{H^{k+1}(\Omega_s)} + \|\boldsymbol{u_d}\|_{H^{k+1}(\Omega_d)} + \|\boldsymbol{u_d}\|^2_{H^{k+1}(\Omega_d)} + \|p_s\|_{H^k(\Omega_s)} + \|p_d\|_{H^k(\Omega_d)} \big),$$

which completes the proof of Theorem 5.1. □

## 6. Numerical Experiments

In this section, we present two numerical results to demonstrate the proposed theoretical results. We explore the phenomenon of convergence-order reduction in different norms. We note that all experiments are tested for $k = 1$ on triangle meshes. The Picard iteration is used to solve the nonlinear system, and the iteration error is set to be $10^{-8}$.

**Example 6.1.** In this example, we consider the problem (2.1)-(2.2) on $\Omega_s = (0, \pi) \times (0, 1)$, $\Omega_d = (0, \pi) \times (-1, 0)$ and the interface be $\Gamma_I = \{0 < x < \pi, y = 0\}$. We set $\mathbf{K}$ to be the identity tensor in $\mathbb{R}^{2\times 2}$, $\mu = \beta = \rho = 1$. The exact solution $\boldsymbol{u} = (\boldsymbol{u}_s, \boldsymbol{u}_d), p = (p_s, p_d)$ are set to be

$$\mathbf{u}_s = \left[ \frac{2}{\pi} \sin(\pi y) \cos(\pi y) \cos(x);\ \left( \left( \frac{\sin(\pi y)}{\pi} \right)^2 - 2 \right) \sin(x) \right], \quad p_s = \sin(x)\sin(y),$$

$$\mathbf{u}_d = -[(e^y - e^{-y}) \cos(x);\ (e^y + e^{-y}) \sin(x)], \qquad p_d = (e^y - e^{-y}) \sin(x).$$

Let the first component of $\boldsymbol{u}$ is $u_1$ and the second component of $\boldsymbol{u}$ is $u_2$. From the error analysis in Theorem 5.1, we expect $\|\boldsymbol{e}_h\|_1 \approx \mathcal{O}(h^1)$, $\|\boldsymbol{e}_{d,0}\|_{L^3(\Omega_d)} \approx \mathcal{O}(h^{2/3})$ and $\|\varepsilon_h\|_h \approx \mathcal{O}(h^1)$. We do not have a theoretical error estimate for $\|\boldsymbol{e}_{s,0}\|_{L^2(\Omega_s)}$ and $\|\boldsymbol{e}_{d,0}\|_{L^2(\Omega_d)}$, but we expect $\|\boldsymbol{e}_{s,0}\|_{L^2(\Omega_s)} \approx \mathcal{O}(h^2)$ and $\|\boldsymbol{e}_{d,0}\|_{L^2(\Omega_d)} \approx \mathcal{O}(h^2)$. The result is shown in Figs. 6.1-6.4 and Table 6.1, where "Iter" represents the number of nonlinear iteration steps. For the sake of brevity in tables, we drop the subscript $L^2$ and $\Omega_l (l = s, d)$ without affecting understanding. In Fig. 6.4, we conveniently denote the norms $\|\boldsymbol{e}_h\|_1$, $\|\boldsymbol{e}_0\|_{L^2(\Omega)}$, $\|\varepsilon_h\|_h$, $\|\boldsymbol{e}_{s,0}\|_{L^2(\Omega_s)}$, $\|\boldsymbol{e}_{d,0}\|_{L^2(\Omega_d)}$, $\|\varepsilon_{s,h}\|_{L^2(\Omega_s)}$ and $\|\varepsilon_{d,h}\|_{L^2(\Omega_d)}$ by "H1-tri", "L2u", "L2p", "L2us", "L2ud", "L2ps" and "L2pd", respectively.

We can see that $\|\boldsymbol{e}_h\|_1 \approx \mathcal{O}(h^1)$ and $\|\varepsilon_h\|_h \approx \mathcal{O}(h^1)$, which are consistent with our theoretical results. However, we note that the order of $\|\varepsilon_{d,h}\|_{L^2(\Omega_d)}$ and $\|\boldsymbol{e}_{d,0}\|_{L^3(\Omega_d)}$ are approximately equal to $\mathcal{O}(h^2)$, which are higher than $\mathcal{O}(h^1)$. The $\|\boldsymbol{e}_{s,0}\|_{L^2(\Omega_s)}$ and $\|\boldsymbol{e}_{d,0}\|_{L^2(\Omega_d)}$ also consistent with our expectation.

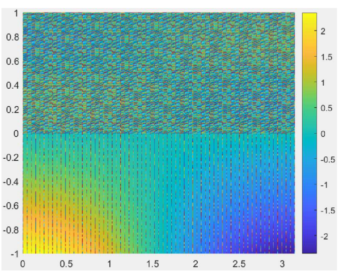

Fig. 6.1. $u_1$ of Example 6.1.

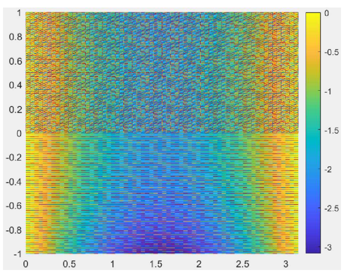

Fig. 6.2. $u_2$ of Example 6.1.

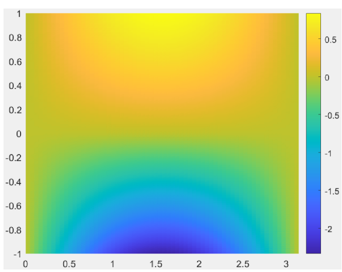

Fig. 6.3. $p$ of Example 6.1.

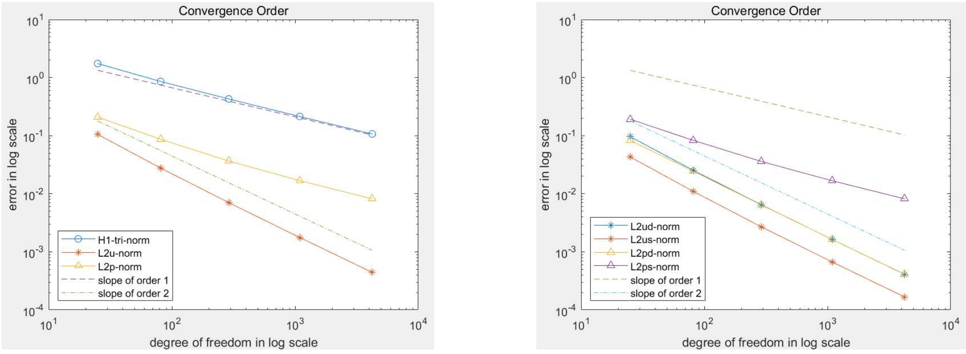


Fig. 6.4. Convergence rates of Example 6.1.

Table 6.1: The errors and the orders of convergence for Example 6.1.

| $h$ | Iter | $\|\boldsymbol{e}_h\|_1$ | Order | $\|\boldsymbol{e_0}\|$ | Order | $\|\varepsilon_h\|_h$ | Order | $\|\boldsymbol{e_{d,0}}\|_{L^3}$ | Order |
|---|---|---|---|---|---|---|---|---|---|
| 4.121e-1 | 11 | 1.746e+0 | | 1.053e-1 | | 2.087e-1 | | 5.292e-2 | |
| 2.061e-1 | 10 | 8.579e-1 | 1.025 | 2.771e-2 | 1.926 | 8.683e-2 | 1.265 | 1.380e-2 | 1.939 |
| 1.030e-1 | 9 | 4.290e-1 | 0.998 | 7.039e-3 | 1.977 | 3.661e-2 | 1.246 | 3.337e-3 | 2.048 |
| 5.151e-2 | 9 | 2.142e-1 | 1.002 | 1.767e-3 | 1.994 | 1.683e-2 | 1.121 | 8.287e-4 | 2.010 |
| 2.576e-2 | 8 | 1.069e-1 | 1.003 | 4.422e-4 | 1.999 | 8.181e-3 | 1.041 | 2.081e-4 | 1.993 |
| $h$ | Iter | $\|\boldsymbol{e_{s,0}}\|$ | Order | $\|\boldsymbol{e_{d,0}}\|$ | Order | $\|\varepsilon_{s,h}\|$ | Order | $\|\varepsilon_{d,h}\|$ | Order |
| 4.121e-1 | 11 | 2.162e-2 | | 6.071e-2 | | 1.144e-1 | | 4.902e-2 | |
| 2.061e-1 | 10 | 5.598e-3 | 1.949 | 1.586e-2 | 1.937 | 3.822e-2 | 1.581 | 1.274e-2 | 1.944 |
| 1.030e-1 | 9 | 1.421e-3 | 1.978 | 4.009e-3 | 1.984 | 1.333e-2 | 1.520 | 3.224e-3 | 1.983 |
| 5.151e-2 | 9 | 3.573e-4 | 1.992 | 1.005e-3 | 1.997 | 5.191e-3 | 1.360 | 8.088e-4 | 1.994 |
| 2.576e-2 | 8 | 8.947e-5 | 1.998 | 2.512e-4 | 1.999 | 2.246e-3 | 1.209 | 2.024e-4 | 1.999 |

**Example 6.2.** In this example, we consider (2.1)-(2.2) on $\Omega_s = (0,1)\times(1,2), \Omega_d = (0,1)\times(0,1)$ and the interface be $\Gamma_I = \{0 < x < 1, y = 1\}$. We also set $\mathbf{K}$ to be the identity tensor in $\mathbb{R}^{2\times 2}$, $\mu = \beta = \rho = 1$. The exact solution $\boldsymbol{u} = (\boldsymbol{u}_s, \boldsymbol{u}_d), p = (p_s, p_d)$ are set to be

$$\mathbf{u}_s = [-\cos(\pi x)\sin(\pi y);\ \ \sin(\pi x)\cos(\pi y)]\,, \quad p_s = \sin(\pi x);$$
$$\mathbf{u}_d = -[\pi(y-1)\cos(\pi x);\ \ \sin(\pi x)], \qquad p_d = y\sin(\pi x).$$

The results are shown in Figs. 6.5-6.8 and Table 6.2. Same as Example 6.1, there is a good agreement between the theoretical results and the numerical results. What is more, $\|\varepsilon_{d,h}\|_{L^2(\Omega_d)}$ and $\|\boldsymbol{e}_{d,0}\|_{L^3(\Omega_d)}$ seem higher than theoretical results.

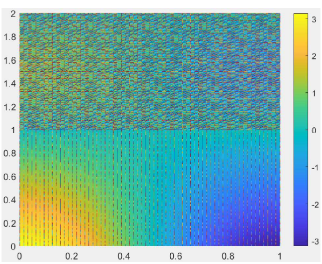

Fig. 6.5. $u_1$ of Example 6.2.

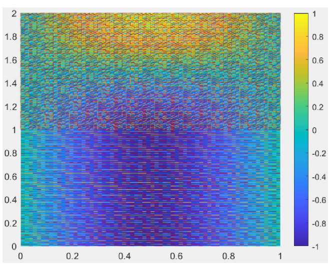

Fig. 6.6. $u_2$ of Example 6.2.

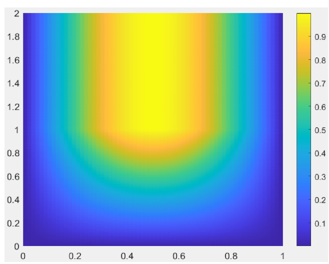

Fig. 6.7. $p$ of Example 6.2.

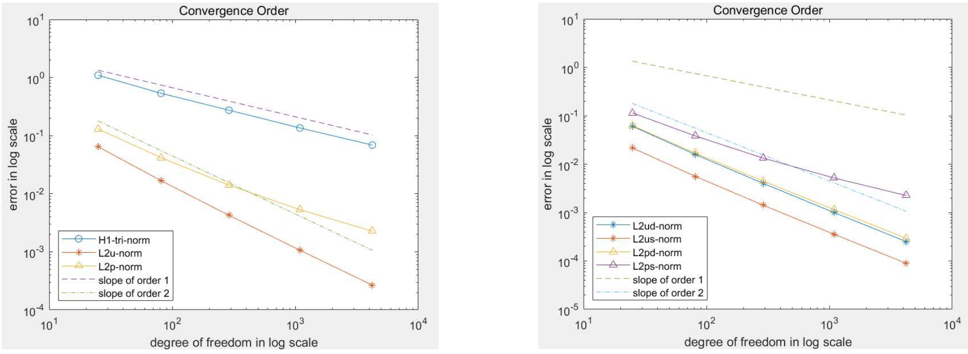


Fig. 6.8. Convergence rates of Example 6.2.

Table 6.2: The errors and the orders of convergence for Example 6.2.

| $h$ | Iter | $\|\!\|\!\|\boldsymbol{e}_h\|\!\|\!\|_1$ | Order | $\|\boldsymbol{e_0}\|$ | Order | $\|\varepsilon_h\|_h$ | Order | $\|\boldsymbol{e_{d,0}}\|_{L^3}$ | Order |
|---|---|---|---|---|---|---|---|---|---|
| 1.768e-1 | 11 | 1.096e+0 | | 6.445e-2 | | 1.303e-1 | | 4.902e-2 | |
| 8.839e-2 | 10 | 5.362e-1 | 1.031 | 1.682e-2 | 1.938 | 4.180e-2 | 1.640 | 1.274e-2 | 1.944 |
| 4.419e-2 | 10 | 2.745e-1 | 0.966 | 4.253e-3 | 1.983 | 1.406e-2 | 1.572 | 3.224e-3 | 1.983 |
| 2.210e-2 | 8 | 1.356e-1 | 1.017 | 1.066e-3 | 1.996 | 5.318e-3 | 1.402 | 8.088e-4 | 1.995 |
| 1.105e-2 | 7 | 6.882e-2 | 0.979 | 2.667e-4 | 1.999 | 2.265e-3 | 1.231 | 2.024e-4 | 1.999 |
| $h$ | Iter | $\|\boldsymbol{e_{s,0}}\|$ | Order | $\|\boldsymbol{e_{d,0}}\|$ | Order | $\|\varepsilon_{s,h}\|$ | Order | $\|\varepsilon_{d,h}\|$ | Order |
| 1.768e-1 | 11 | 2.162e-2 | | 6.071e-2 | | 1.144e-1 | | 6.240e-2 | |
| 8.839e-2 | 10 | 5.598e-3 | 1.949 | 1.586e-2 | 1.937 | 3.822e-2 | 1.581 | 1.962e-2 | 1.883 |
| 4.419e-2 | 10 | 1.421e-3 | 1.978 | 4.009e-3 | 1.984 | 1.333e-2 | 1.520 | 4.469e-3 | 1.921 |
| 2.210e-2 | 8 | 3.573e-4 | 1.992 | 1.005e-3 | 1.997 | 5.191e-3 | 1.360 | 1.153e-3 | 1.955 |
| 1.105e-2 | 7 | 8.947e-5 | 1.998 | 2.512e-4 | 1.999 | 2.246e-3 | 1.209 | 2.930e-4 | 1.976 |

## 7. Conclusion

We have applied the WG method to solve the coupled Stokes and Darcy-Forchheimer model. We established the WG scheme and proved the existence and uniqueness of the scheme by the discrete inf-sup condition and Minty-Browder theorem. When the exact solution $\boldsymbol{u} \in [H^{k+1}(\Omega_s)]^2 \times [H^{k+1}(\Omega_d)]^2$ and $p \in H^k(\Omega_s) \times H^k(\Omega_d)$, we prove that the orders of convergence for $|\!|\!|\boldsymbol{e}_h|\!|\!|_1, \|\varepsilon_h\|_h$, and $\|\boldsymbol{e}_{d,0}\|_{L^2(\Omega_d)}$ are $\mathcal{O}(h^{\min\{k,3k/4+3/4\}})$, and the order of convergence for $\|\boldsymbol{e}_{d,0}\|_{L^3(\Omega_d)}$ is $\mathcal{O}(h^{\min\{2k/3,k/2+1/2\}})$. The numerical results for $\boldsymbol{u}$ and $p$ are optimal, which are in agreement with our theoretical analysis.

The main difficulty in the theoretical analysis of WG method is to derive the orders of error estimates when the true solution in $[H^{k+1}]^2$ and $H^k$ space. From our numerical results, when $k = 1$, the order of $\|\varepsilon_{d,h}\|_{L^2(\Omega_d)}$ are approximately $\mathcal{O}(h^2)$, which is higher than the results $\mathcal{O}(h^1)$ in [1, 42, 45]. There are also further questions worth analyzing:

- extension of the results to 3D problems,
- derivation of the convergence estimates of $\|\boldsymbol{e}_{s,0}\|_{L^2(\Omega_s)}$,
- further analysis of the influence of the nonlinearity on the order of convergence of the method.

## Appendix A

In the appendix, we present some technical preparations for the proof of the existence and uniqueness of the numerical solution and the optimal error estimate.

**Lemma A.1 ([10]).** *The projection operators $Q_h, \mathbb{Q}_h$ and $\pi_h$ satisfy*

$$\nabla_w(Q_h\boldsymbol{w}) = \mathbb{Q}_h(\nabla\boldsymbol{w}), \qquad \forall\, \boldsymbol{w} \in [H^1(\Omega)]^2, \tag{A.1}$$

$$\nabla_w \cdot (Q_h\boldsymbol{w}) = \pi_h(\nabla\cdot\boldsymbol{w}), \quad \forall\, \boldsymbol{w} \in H(\mathrm{div},\Omega). \tag{A.2}$$

On the regular polygonal partition, from [37], we know the following trace inequality and inverse inequality also holds.

**Lemma A.2 (Trace Inequality, [24]).** *Assume that the partition $\mathcal{T}_h$ is shape-regular. Suppose $\varphi \in H^1(T)$, then there exists a constant $C$ such that the following inequality holds on each cell $T \in \mathcal{T}_h$:*

$$\|\varphi\|^2_{L^2(\partial T)} \le C\big(h_T^{-1}\|\varphi\|^2_{L^2(T)} + h_T\|\nabla\varphi\|^2_{L^2(T)}\big). \tag{A.3}$$

**Lemma A.3 (Inverse Inequality, [43]).** *Assume that the partition $\mathcal{T}_h$ is shape-regular. Suppose $\psi \in P_k(T)$, then there exists a constant $C$ such that on each cell $T \in \mathcal{T}_h$,*

$$\|\nabla\psi\|_{L^2(T)} \le Ch_T^{-1}\|\psi\|_{L^2(T)}. \tag{A.4}$$

The vector version of the trace inequality and the inverse inequality are trivial.

**Lemma A.4 ([38]).** *Let $\mathcal{T}_h$ be a polygonal partition of $\Omega$ satisfying the shape-regular assumptions. $\boldsymbol{\phi} \in [H^{r+1}(\Omega)]^2$ and $q \in H^r(\Omega)$ with $0 \le r \le k$. Then, On each cell $T \in \mathcal{T}_h$, for $0 \le s \le 1$, there exists a constant $C$ such that*

$$\sum_{T\in\mathcal{T}_h} h_T^{2s}\|\boldsymbol{\phi} - Q_0\boldsymbol{\phi}\|^2_{H^s(T)} \le Ch^{2(r+1)}\|\boldsymbol{\phi}\|^2_{H^{r+1}(\Omega)}, \tag{A.5}$$

$$\sum_{T\in\mathcal{T}_h} h_T^{2s}\|\nabla\boldsymbol{\phi} - \mathbb{Q}_h(\nabla\boldsymbol{\phi})\|^2_{H^s(T)} \le Ch^{2r}\|\boldsymbol{\phi}\|^2_{H^{r+1}(\Omega)}, \tag{A.6}$$

$$\sum_{T\in\mathcal{T}_h} h_T^{2s}\|q - \pi_h q\|^2_{H^s(T)} \le Ch^{2r}\|p\|^2_{H^r(\Omega)}, \tag{A.7}$$

$$\sum_{T\in\mathcal{T}_h} \|\boldsymbol{\phi} - Q_b\boldsymbol{\phi}\|^2_{L^2(\partial T)} \le Ch^{2(r+1)}\|\boldsymbol{\phi}\|^2_{H^{r+1}(\partial\Omega)}. \tag{A.8}$$

**Lemma A.5 ([11]).** *Suppose $m$ and $l$ are non-negative integers, $\Omega$ is a smooth domain in $\mathbb{R}^n$. When $mp = n$, we have*

$$W^{m,p}(\Omega) \ \hookrightarrow\ W^{m-\frac{1}{p},p}(\partial\Omega) \ \hookrightarrow\ L^q(\partial\Omega), \quad \forall\, 1 \le q < \infty,$$

*and if $0 \le l \le m$, then*

$$W^{m,p}(\Omega) \ \hookrightarrow\ W^{l,p}(\Omega).$$

**Lemma A.6 (Young's Inequality).** *For $a, b \in R$ and $p, q > 1$ satisfying $1/p + 1/q = 1$, we have*

$$ab \le \epsilon a^p + C(\epsilon)b^q,$$

*where $\epsilon$ is sufficiently small and $C(\epsilon) = (p\epsilon)^{-q/p}/q$.*

**Lemma A.7 ([8]).** *Assume that the partition $\mathcal{T}_{d,h}$ is shape-regular. Then for each $p,q \in [1,\infty)$, there exists a constant $C>0$ independent of $h_{T_d}$ such that for any $T_d \in \mathcal{T}_{d,h}$,*

$$h_{T_d}^{-\frac{2}{p}} \|\boldsymbol{v}_{d,0}\|_{L^p(\Omega_d)} \leq C h_{T_d}^{-\frac{2}{q}} \|\boldsymbol{v}_{d,0}\|_{L^q(\Omega_d)}, \quad \forall\, \boldsymbol{v}_{d,h} = \{\boldsymbol{v}_{d,0}, \boldsymbol{v}_{d,b}\} \in V_{d,h}. \tag{A.9}$$

**Acknowledgments.** We sincerely thank the anonymous reviewers for their insightful comments, which have helped improve the quality of this paper.

H. Peng was supported by the National Natural Science Foundation of China (NSFC) (Grant No. 12301519). Q. Zhai was partially supported by the National Natural Science Foundation of China (Grant Nos. 12271208, 11901015).

## References

[1] J.A. Almonacid, H.S. Díaz, G.N. Gatica, and A. Márquez, A fully mixed finite element method for the coupling of the Stokes and Darcy-Forchheimer problems, *IMA J. Numer. Anal.*, **40**:2 (2020), 1454–1502.

[2] T. Arbogast and D.S. Brunson, A computational method for approximating a Darcy-Stokes system governing a vuggy porous medium, *Comput. Geosci.*, **11**:3 (2007), 207–218.

[3] T. Arbogast and M.S.M. Gomez, A discretization and multigrid solver for a Darcy-Stokes system of three dimensional vuggy porous media, *Comput. Geosci.*, **13**:3 (2009), 331–348.

[4] T. Arbogast and H.L. Lehr, Homogenization of a Darcy-Stokes system modeling vuggy porous media, *Comput. Geosci.*, **10**:3 (2006), 291–302.

[5] J.D. Audu, F. Fairag, and K. Mustapha, Mixed finite element analysis for generalized Darcy-Forchheimer model in porous media, *J. Comput. Appl. Math.*, **353** (2019), 191–203.

[6] J. Bear, *Dynamics of Fluids in Porous Media*, American Elsevier, 1972.

[7] F. Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers, *Rev. Française Autom. Informat. Rech. Opérationnelle Sér.*, **8**:2 (1974), 129–151.

[8] A. Buffa and C. Ortner, Compact embeddings of broken Sobolev spaces and applications, *IMA J. Numer. Anal.*, **29**:4 (2009), 827–855.

[9] Y.Z. Cao, M. Gunzburger, X.L. Hu, F. Hua, X.M. Wang, and W.D. Zhao, Finite element approximations for Stokes-Darcy flow with Beavers-Joseph interface conditions, *SIAM J. Numer. Anal.*, **47**:6 (2010), 4239–4256.

[10] W.B. Chen, F. Wang, and Y.Q. Wang, Weak Galerkin method for the coupled Darcy-Stokes flow, *IMA J. Numer. Anal.*, **36**:2 (2016), 897–921.

[11] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, SIAM, 1978.

[12] P.G. Ciarlet, *Linear and Nonlinear Functional Analysis with Applications*, SIAM, 2008.

[13] M. Discacciati, E. Miglio, and A. Quarteroni, Mathematical and numerical models for coupling surface and groundwater flows, *Appl. Numer. Math.*, **43**:1-2 (2002), 57–74.

[14] M. Discacciati, A. Quarteroni, and A. Valli, Robin-Robin domain decomposition methods for the Stokes-Darcy coupling, *SIAM J. Numer. Anal.*, **45**:3 (2007), 1246–1268.

[15] N.S. Hanspal, A.N. Waghode, V. Nassehi, and R.J. Wakeman, Numerical analysis of coupled Stokes/Darcy flows in industrial filtrations, *Transp. Porous Media*, **64**:1 (2006), 73–101.

[16] I. Igreja and A.F.D. Loula, A stabilized hybrid mixed DGFEM naturally coupling Stokes-Darcy flows, *Comput. Methods Appl. Mech. Eng.*, **339** (2018), 739–768.

[17] W. Layton, H. Tran, and C. Trenchea, Analysis of long time stability and errors of two partitioned methods for uncoupling evolutionary groundwater-surface water flows, *SIAM J. Numer. Anal.*, **51**:1 (2013), 248–272.

[18] W.J. Layton, F. Schieweck, and I. Yotov, Coupling fluid flow with porous media flow, *SIAM J. Numer. Anal.*, **40**:6 (2002), 2195–2218.

[19] M.H. Li, D.Y. Shi, Z.Z. Li, and H.R. Chen, Two-level mixed finite element methods for the Navier-Stokes equations with damping, *J. Math. Anal. Appl.*, **470**:1 (2019), 292–307.

[20] R. Li, Y.L. Gao, J. Li, and Z.X. Chen, A weak Galerkin finite element method for a coupled Stokes-Darcy problem on general meshes, *J. Comput. Appl. Math.*, **334** (2018), 111–127.

[21] G. Lin, J.G. Liu, L. Mu, and X. Ye, Weak Galerkin finite element methods for Darcy flow: Anisotropy and heterogeneity, *J. Comput. Phys.*, **276** (2014), 422–437.

[22] J.G. Liu, S. Tavener, and Z.R. Wang, Lowest-order weak Galerkin finite element method for Darcy flow on convex polygonal meshes, *SIAM J. Sci. Comput.*, **40**:5 (2018), B1229–B1252.

[23] L. Mu, J.P. Wang, and X. Ye, Weak Galerkin finite element methods on polytopal meshes, *Int. J. Numer. Anal. Model.*, **12** (2015), 31–53.

[24] L. Mu, J.P. Wang, and X. Ye, A weak Galerkin finite element method with polynomial reduction, *J. Comput. Appl. Math.*, **285** (2013), 45–58.

[25] H. Pan and H.X. Rui, Mixed element method for two-dimensional Darcy-Forchheimer model, *J. Sci. Comput.*, **52**:3 (2012), 563–587.

[26] H. Peng, Q.L. Zhai, R. Zhang, and S.Y. Zhang, A weak Galerkin-mixed finite element method for the Stokes-Darcy problem, *Sci. China Math.*, **64**:10 (2021), 2357–2380.

[27] M.Z. Qin, R.S. Wang, Q.L. Zhai, and R. Zhang, Weak Galerkin method for second-order elliptic equations with Newton boundary condition, *Commun. Comput. Phys.*, **33**:2 (2023), 568–595.

[28] B. Riviére, Analysis of a discontinuous finite element method for the coupled Stokes and Darcy problems, *J. Sci. Comput.*, **22-23**:1-3 (2005), 479–500.

[29] B. Rivière and I. Yotov, Locally conservative coupling of Stokes and Darcy flows, *SIAM J. Numer. Anal.*, **42**:5 (2005), 1959–1977.

[30] H.X. Rui and W. Liu, A two-grid block-centered finite difference method for Darcy-Forchheimer flow in porous media, *SIAM J. Numer. Anal.*, **53**:4 (2015), 1941–1962.

[31] H.X. Rui and H. Pan, A block-centered finite difference method for slightly compressible Darcy-Forchheimer flow in porous media, *J. Sci. Comput.*, **73**:1 (2017), 70–92.

[32] H.X. Rui and J.Y. Zhang, A stabilized mixed finite element method for coupled Stokes and Darcy flows with transport, *Comput. Methods Appl. Mech. Eng.*, **315** (2017), 169–189.

[33] H.X. Rui, D.H. Zhao, and H. Pan, A block-centered finite difference method for Darcy-Forchheimer model with variable Forchheimer number, *Numer. Methods Partial Differential Equations*, **31**:5 (2015), 1603–1622.

[34] P.G. Saffman, On the boundary condition at the surface of a porous medium, *Stud. Appl. Math.*, **50**:2 (1971), 93–101.

[35] C.M. Wang and J.P. Wang, Weak Galerkin finite element methods for elliptic PDEs, *Sci. Sin. Math.*, **45**:7 (2015), 1061–1092.

[36] J.P. Wang and X. Ye, A weak Galerkin finite element method for second-order elliptic problems, *J. Comput. Appl. Math.*, **241** (2013), 103–115.

[37] J.P. Wang and X. Ye, A weak Galerkin mixed finite element method for second order elliptic problems, *Math. Comp.*, **289** (2014), 2101–2126.

[38] J.P. Wang and X. Ye, A weak Galerkin finite element method for the Stokes equations, *Adv. Comput. Math.*, **42**:1 (2016), 155–174.

[39] J.P. Wang, X. Ye, Q.L. Zhai, and R. Zhang, Discrete maximum principle for the $P_1$-$P_0$ weak Galerkin finite element approximations, *J. Comput. Phys.*, **362** (2018), 114–130.

[40] Q.L. Zhai, H.H. Xie, R. Zhang, and Z.M. Zhang, The weak Galerkin method for elliptic eigenvalue problems, *Commun. Comput. Phys.*, **26**:1 (2019), 160–191.

[41] Q.L. Zhai, R. Zhang, and L. Mu, A new weak Galerkin finite element scheme for the Brinkman model, *Commun. Comput. Phys.*, **19** (2016), 1409–1434.

[42] J.Y. Zhang and H.X. Rui, A stabilized Crouzeix-Raviart element method for coupling Stokes and Darcy-Forchheimer flows, *Numer. Methods Partial Differential Equations*, **33**:4 (2017), 1070–1094.

[43] Q. Zhang and Z. Ran, A weak Galerkin mixed finite element method for second-order elliptic

equations with Robin boundary conditions, *J. Comput. Math.*, **34**:5 (2016), 532–548.

[44] R. Zhang and Q.L. Zhai, A weak Galerkin finite element scheme for the biharmonic equations by using polynomials of reduced order, *J. Sci. Comput.*, **64**:2 (2015), 559–585.

[45] L. Zhao, E.T. Chung, E.-J. Park, and G. Zhou, Staggered DG method for coupling of the Stokes and Darcy-Forchheimer problems, *SIAM J. Numer. Anal.*, **59**:1 (2021), 1–31.