Journal of Computational Mathematics
Vol.xx, No.x, 2023, 1–26.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2211-m2021-0255

MIXED DISCONTINUOUS GALERKIN METHOD FOR QUASI-NEWTONIAN STOKES FLOWS*

Yanxia Qian
School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China
Email: yxqian0520@163.com
Fei Wang[1)]
School of Mathematics and Statistics & State Key Laboratory of Multiphase Flow in Power Engineering, Xi'an Jiaotong University, Xi'an 710049, China
Email: feiwang.xjtu@xjtu.edu.cn
Wenjing Yan
School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China
Email: wenjingyan@xjtu.edu.cn

Abstract

In this paper, we introduce and analyze an augmented mixed discontinuous Galerkin (MDG) method for a class of quasi-Newtonian Stokes flows. In the mixed formulation, the unknowns are strain rate, stress and velocity, which are approximated by a discontinuous piecewise polynomial triplet $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$ for $k \geq 0$. Here, the discontinuous piecewise polynomial function spaces for the field of strain rate and the stress field are designed to be symmetric. In addition, the pressure is easily recovered through simple postprocessing. For the benefit of the analysis, we enrich the MDG scheme with the constitutive equation relating the stress and the strain rate, so that the well-posedness of the augmented formulation is obtained by a nonlinear functional analysis. For $k \geq 0$, we get the optimal convergence order for the stress in broken $\underline{\boldsymbol{H}}(\mathbf{div})$-norm and velocity in $\boldsymbol{L}^2$-norm. Furthermore, the error estimates of the strain rate and the stress in $\underline{\boldsymbol{L}}^2$-norm, and the pressure in L^2-norm are optimal under certain conditions. Finally, several numerical examples are given to show the performance of the augmented MDG method and verify the theoretical results. Numerical evidence is provided to show that the orders of convergence are sharp.

Mathematics subject classification: 65N30, 65M60.
Key words: Quasi-Newtonian flows, Mixed discontinuous Galerkin method, Symmetric strain rate, Symmetric stress, Optimal convergence orders.

1. Introduction

The quasi-Newtonian Stokes equations arise in modeling flows of biological fluids, lubricants, paints, polymeric fluids, where the fluid viscosity is assumed to be a nonlinear function of the strain rate tensor [30, 35]. Let Ω be a bounded and simply connected polygonal domain in $\mathbb{R}^n$ with Lipschitz continuous boundary Γ. In this paper, we consider a class of Stokes equations whose viscosity depends nonlinearly on the strain rate, which is a characteristic feature of quasi-Newtonian flows: Given $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{g} \in \boldsymbol{H}^{1/2}(\Gamma)$, find a stress field $\underline{\boldsymbol{\sigma}}$, a velocity field $\boldsymbol{u}$

* Received September 3, 2021 / Revised version received April 6, 2022 / Accepted November 7, 2022 / Published online June 5, 2023 /
1) Corresponding author

and a pressure field p such that

$$\underline{\boldsymbol{\sigma}} = 2\mu(|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|)\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}) - p\underline{\boldsymbol{I}} \quad \text{in } \Omega, \tag{1.1a}$$

$$\mathbf{div}\underline{\boldsymbol{\sigma}} = -\boldsymbol{f} \quad \text{in } \Omega, \tag{1.1b}$$

$$\mathrm{div}\boldsymbol{u} = 0 \quad \text{in } \Omega, \tag{1.1c}$$

$$\boldsymbol{u} = \boldsymbol{g} \quad \text{on } \Gamma, \tag{1.1d}$$

$$\int_\Omega p d\boldsymbol{x} = 0, \tag{1.1e}$$

where $\mu : \mathbb{R}^+ \to \mathbb{R}^+$ denotes the nonlinear kinematic viscosity function of the fluid and $|\cdot|$ stands for Euclidean norm of tensors in $\mathbb{R}^{n\times n}$. Due to the incompressibility condition, we assume that $\boldsymbol{g}$ satisfies the compatibility condition $\int_\Gamma \boldsymbol{g}\cdot\boldsymbol{n}ds = 0$, where $\boldsymbol{n}$ stands for the unit outward normal on Γ. Let us list some classic examples of the nonlinear kinematic viscosity μ for the quasi-Newtonian flows.

Power law:

$$\mu(t) = \mu_0 t^{\beta-2}, \quad \forall t \in \mathbb{R}^+ \quad \text{with} \quad \mu_0 > 0 \quad \text{and} \quad 1 < \beta < 2$$

serves to model the viscosity of many polymeric solutions and melts over a considerable range of shear rates [30].

Ladyzhenskaya law:

$$\mu(t) = (\mu_0 + \mu_1 t)^{\beta-2}, \quad \forall t \in \mathbb{R}^+ \quad \text{with} \quad \mu_0 \geq 0, \quad \mu_1 > 0 \quad \text{and} \quad \beta > 1$$

is used to model the fluids with large stresses [35].

Carreau law:

$$\mu(t) = \mu_0 + \mu_1(1+t^2)^{\frac{\beta-2}{2}}, \quad \forall t \in \mathbb{R}^+ \quad \text{with} \quad \mu_0 \geq 0, \quad \mu_1 > 0 \quad \text{and} \quad \beta \geq 1$$

is applied to model visco-plastic flows and creeping flow of metals [37].

The linear Stokes problem is recovered from the above laws when $\beta = 2$.

Many researchers aim at studying the efficient numerical methods for quasi-Newtonian flows and related problems, such as the conforming and nonconforming finite element method [4,6,17], the mixed finite element method [5, 21, 22], the dual-mixed finite element method [18, 31], the discontinuous Galerkin (DG) method [13, 16, 26, 27], the weak Galerkin method [43] and the virtual element method [14, 25] and so on. Traditionally, the numerical methods are studied based on the velocity-pressure variational formulation, where the velocity and the pressure are the main unknowns [8, 11]. Over the past decades, many researchers have paid attention to stress-based and pseudostress-based formulations [7, 15, 19, 23] because they provide a unified framework for both the Newtonian and non-Newtonian flows. Actually, a formulation comprising the stress as a fundamental unknown is unavoidable for non-Newtonian flows in which the constitutive law is nonlinear. Therefore, the mixed formulation is a good choice; besides the original unknowns, it yields direct approximations of several other physical interest quantities. For example, it is very desirable to calculate stress accurately and directly for flow problems involving interaction with solid structures. In [26], a pseudostress-based hybrid DG (HDG) scheme with BDM-like (Brezzi-Douglas-Marini) elements for the quasi-Newtonian Stokes flows was studied, and a priori error analysis was given. However, the $\underline{\boldsymbol{L}}^2$ error estimates of strain rate and stress are not optimal. Then, the HDG scheme with an RT-like (Raviart-Thomas) elements was studied in [27], and optimal error analysis is established; in addition, a reliable and efficient residual-based a posteriori error estimator was derived. Note that these methods are based on pseudostress formulation, in which the pseudostress tensor is not symmetric. According to the principle of conservation of angular momentum, the stress tensor needs to be symmetric, but it is very challenging to develop such stable mixed finite element methods [1]. We refer the readers to [1, 29, 32–34, 42] and the references therein for this topic.

In this paper, we aim at constructing an augmented mixed discontinuous Galerkin method with symmetric strain rate and stress for solving the quasi-Newtonian Stokes flows in the stress-strain-velocity-based formulation. In order to deal with the aforementioned nonlinearity, we introduce the strain rate as a new unknown [21]. Then, we choose a discontinuous piecewise polynomial triplet $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$ $(k \geq 0)$ to approximate the strain rate, stress and velocity, respectively. Here, the superscript $\mathbb{S}$ means that the field of strain rate and the stress field are approximated by symmetric polynomials. Moreover, we modify the variational formulation by adding a redundant equation arising from the constitutive law relating the stress and the strain rate, which allows us to prove the well-posedness of the formulation by nonlinear functional analysis. The main results of this article include that:

(i) The well-posedness of both continuous and discrete schemes are obtained.

(ii) For $k \geq 0$, we get the optimal convergence order for the stress in broken $\underline{\boldsymbol{H}}(\mathbf{div})$-norm and velocity in $\boldsymbol{L}^2$-norm.

(iii) For $k \geq n$, the optimal $\underline{\boldsymbol{L}}^2$ error estimates for the strain rate and the stress are obtained, and then, the optimal L^2 error estimate for pressure can be derived via a postprocessing calculation.

The rest of the paper is organized as follows. In Section 2, we introduce the stress-strain-velocity-based formulation for the quasi-Newtonian Stokes flows and present some preliminary results. In Section 3, the MDG scheme is introduced, and the well-posedness is obtained. We show the stability of the discrete scheme and prove optimal error estimates for all variables in Section 4. In Section 5, numerical examples are provided to confirm the theoretical results and to illustrate the performance of the mixed DG scheme. Finally, we give a short summary in Section 6.

2. Mathematical Setting of the Continuous Problem

In this section, the stress-strain-velocity formulation of the quasi-Newtonian Stokes flows is introduced and the stability analysis of the continuous problem is provided.

2.1. Notation

Given an integer $m \geq 0$ and a bounded subdomain $D \subset \mathbb{R}^n,\ n = 2, 3$, we denote the scalar-valued Sobolev space by $H^m(D) = W^{m,2}(D)$ with the norm $\|\cdot\|_{m,D}$ and seminorm $|\cdot|_{m,D}$. As $m = 0, H^0(D)$ coincides with the Lebesgue spaces $L^2(D)$, equipped with the usual L^2-inner product $(\cdot,\cdot)_D$ and L^2-norm $\|\cdot\|_{0,D}$. The L^2-inner product (or duality pairing) on ∂D is denoted by $\langle\cdot,\cdot\rangle_{\partial D}$. If D is chosen as Ω, we abbreviate $(\cdot,\cdot)_\Omega$ and $\langle\cdot,\cdot\rangle_{\partial\Omega}$ by using $(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle$, respectively. Similar rule follows for the norms defined later. We denote the vector-valued function spaces, tensor-valued function spaces and symmetric tensor-valued function spaces, whose entries are in $H^m(D)$, by $\boldsymbol{H}^m(D)$, $\underline{\boldsymbol{H}}^m(D)$ and $\underline{\boldsymbol{H}}^m(D;\mathbb{S})$, respectively. Here, $\mathbb{S}$ denotes the real symmetric matrix order $n\times n$. In particular, $\boldsymbol{H}^0(D)=\boldsymbol{L}^2(D)$, $\underline{\boldsymbol{H}}^0(D)=\underline{\boldsymbol{L}}^2(D)$ and $\underline{\boldsymbol{H}}^0(D;\mathbb{S})=\underline{\boldsymbol{L}}^2(D;\mathbb{S})$. For matrices

$$\underline{\boldsymbol{\tau}}=(\tau_{ij})\in\mathbb{R}^{n\times n},\quad \underline{\boldsymbol{\zeta}}=(\zeta_{ij})\in\mathbb{R}^{n\times n},$$

we write as usual

$$\underline{\boldsymbol{\tau}}^t=(\tau_{ji}),\quad \operatorname{tr}(\underline{\boldsymbol{\tau}})=\sum_{i=1}^n\tau_{ii},\quad \underline{\boldsymbol{\tau}}^d=\underline{\boldsymbol{\tau}}-\frac{1}{n}\operatorname{tr}(\underline{\boldsymbol{\tau}})\underline{\boldsymbol{I}},\quad \underline{\boldsymbol{\tau}}:\underline{\boldsymbol{\zeta}}=\sum_{i,j=1}^n\tau_{ij}\zeta_{ij}, \tag{2.1}$$

where $\underline{\boldsymbol{I}}$ is the identity matrix. Then, we introduce the following space:

$$\underline{\boldsymbol{H}}(\mathbf{div},D)=\big\{\underline{\boldsymbol{\tau}}\in\underline{\boldsymbol{L}}^2(D):\mathbf{div}\underline{\boldsymbol{\tau}}\in\boldsymbol{L}^2(D)\big\},$$

equipped with the norm

$$\|\underline{\boldsymbol{\tau}}\|_{\mathbf{div},D}=\big(\|\underline{\boldsymbol{\tau}}\|_{0,D}^2+\|\mathbf{div}\underline{\boldsymbol{\tau}}\|_{0,D}^2\big)^{\frac{1}{2}}.$$

Here, $\mathbf{div}$ stands for the usual divergence operator div acting along each row of tensor, i.e., the i-th row of $\mathbf{div}\underline{\boldsymbol{\sigma}}$ is the divergence of the i-th row vector of the matrix $\underline{\boldsymbol{\sigma}}$. And the i-th row of the matrix $\underline{\boldsymbol{\nabla}}(\boldsymbol{u})$ in the strain rate tensor

$$\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})=\frac{1}{2}\big(\underline{\boldsymbol{\nabla}}(\boldsymbol{u})+\underline{\boldsymbol{\nabla}}(\boldsymbol{u})^t\big)$$

is the gradient (written as a row) of the i-th component of the vector $\boldsymbol{u}$. Similarly, we define the symmetric $\underline{\boldsymbol{H}}(\mathbf{div},D;\mathbb{S})$ function space. As mentioned above, we simplify the spaces $\underline{\boldsymbol{H}}(\mathbf{div},\Omega)$ and $\underline{\boldsymbol{H}}(\mathbf{div},\Omega;\mathbb{S})$ to $\underline{\boldsymbol{H}}(\mathbf{div})$ and $\underline{\boldsymbol{H}}(\mathbf{div};\mathbb{S})$, respectively.

2.2. Augmented stress-strain-velocity-based formulation

Let $\psi_{ij}:\mathbb{R}^{n\times n}\to\mathbb{R}$ be a mapping given by $\psi_{ij}(\underline{\boldsymbol{s}})=\mu(|\underline{\boldsymbol{s}}|)s_{ij}$ for all $\underline{\boldsymbol{s}}=(s_{ij})\in\mathbb{R}^{n\times n}$, $i,j\in\{1,\dots,n\}$. In order to prove the strong monotonicity and Lipschitz-continuity properties of the continuous and discrete nonlinear operators involving the viscosity function μ, we assume that μ is of class C^1 and there exist $r_0,r_1>0$ such that for all $\underline{\boldsymbol{\theta}}=(\theta_{ij})\in\mathbb{R}^{n\times n}$ and $\underline{\boldsymbol{s}}=(s_{ij})\in\mathbb{R}^{n\times n}$,

$$\sum_{i,j,k,l=1}^n\frac{\partial}{\partial\underline{\boldsymbol{s}}_{kl}}\psi_{ij}(\underline{\boldsymbol{s}})\theta_{ij}\theta_{kl}\geq r_0\|\underline{\boldsymbol{\theta}}\|_0^2, \tag{2.2}$$

and

$$|\psi_{ij}(\underline{\boldsymbol{s}})|\leq r_1\|\underline{\boldsymbol{s}}\|_0,\quad \left|\frac{\partial}{\partial\underline{\boldsymbol{s}}_{kl}}\psi_{ij}(\underline{\boldsymbol{s}})\right|\leq r_1,\quad \forall\, i,j,k,l\in\{1,\dots,n\}. \tag{2.3}$$

The Carreau law satisfies (2.2) and (2.3) for all $\mu_0>0,\mu_1>0$ and $1\leq\beta\leq 2$ [37,39].

From the Eqs. (1.1a) and (1.1c), we observe that

$$p=-\frac{1}{n}\operatorname{tr}(\underline{\boldsymbol{\sigma}})\quad\text{in }\ \Omega. \tag{2.4}$$

Then, introducing an auxiliary unknown $\underline{\boldsymbol{t}} = \underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})$ and eliminating the pressure in the problem (1.1), we can rewrite it as [20, 25]

$$\begin{aligned}
&\underline{\boldsymbol{\sigma}}^d = 2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}) && \text{in } \Omega, && (2.5a)\\
&\underline{\boldsymbol{t}} = \underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}) && \text{in } \Omega, && (2.5b)\\
&\mathbf{div}\underline{\boldsymbol{\sigma}} = -\boldsymbol{f} && \text{in } \Omega, && (2.5c)\\
&\mathrm{tr}(\underline{\boldsymbol{t}}) = 0 && \text{in } \Omega, && (2.5d)\\
&\boldsymbol{u} = \boldsymbol{g} && \text{on } \Gamma, && (2.5e)\\
&\int_\Omega \mathrm{tr}(\underline{\boldsymbol{\sigma}}) d\boldsymbol{x} = 0, && && (2.5f)
\end{aligned}$$

where $\underline{\boldsymbol{\psi}} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is given by $\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{s}}) = (\psi_{ij}(\underline{\boldsymbol{s}})) = (\mu(|\underline{\boldsymbol{s}}|)s_{ij})$ for all $\underline{\boldsymbol{s}} = (s_{ij}) \in \mathbb{R}^{n\times n}$.

Set

$$\begin{aligned}
&\underline{\boldsymbol{\Sigma}}_1 = \{\underline{\boldsymbol{s}} \in \underline{\boldsymbol{L}}^2(\Omega;\mathbb{S}) : \mathrm{tr}(\underline{\boldsymbol{s}}) = 0\},\\
&\underline{\boldsymbol{\Sigma}}_2 = \left\{\underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{H}}(\mathbf{div};\mathbb{S}) : \int_\Omega \mathrm{tr}(\underline{\boldsymbol{\tau}}) d\boldsymbol{x} = 0\right\},\\
&\boldsymbol{V} = \boldsymbol{L}^2(\Omega).
\end{aligned}$$

Due to the fact that $\|\underline{\boldsymbol{\tau}}\|_0^2 \lesssim \|\underline{\boldsymbol{\tau}}^d\|_0^2 + \|\mathbf{div}\underline{\boldsymbol{\tau}}\|_0^2$ (see [3, Lemma 2.3]) for all $\underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{\Sigma}}_2$, $\underline{\boldsymbol{\Sigma}}_2$ is endowed with the norm

$$\|\underline{\boldsymbol{\tau}}\|_{\underline{\boldsymbol{\Sigma}}_2}^2 = \left(\underline{\boldsymbol{\tau}}^d, \underline{\boldsymbol{\tau}}^d\right) + (\mathbf{div}\underline{\boldsymbol{\tau}}, \mathbf{div}\underline{\boldsymbol{\tau}}), \quad \forall\, \underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{\Sigma}}_2. \tag{2.6}$$

Furthermore, by the definition of $\underline{\boldsymbol{\tau}}^d$, it is easy to check that

$$\|\underline{\boldsymbol{\tau}}\|_0^2 = \|\underline{\boldsymbol{\tau}}^d\|_0^2 + \frac{1}{n}\|\mathrm{tr}(\underline{\boldsymbol{\tau}})\|_0^2, \quad \|\mathrm{tr}(\underline{\boldsymbol{\tau}})\|_0 \le \sqrt{n}\|\underline{\boldsymbol{\tau}}\|_0. \tag{2.7}$$

The variational formulation of problem (2.5) reads as follows: Given $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{g} \in \boldsymbol{H}^{1/2}(\Gamma)$, find $(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}, \boldsymbol{u}) \in \underline{\boldsymbol{\Sigma}}_1 \times \underline{\boldsymbol{\Sigma}}_2 \times \boldsymbol{V}$ such that

$$\begin{aligned}
&\left(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}), \underline{\boldsymbol{s}}\right) - (\underline{\boldsymbol{\sigma}}^d, \underline{\boldsymbol{s}}) = 0, && \forall\, \underline{\boldsymbol{s}} \in \underline{\boldsymbol{\Sigma}}_1, && (2.8a)\\
&(\underline{\boldsymbol{\tau}}^d, \underline{\boldsymbol{t}}) + (\mathbf{div}\underline{\boldsymbol{\tau}}, \boldsymbol{u}) = \langle \underline{\boldsymbol{\tau}}\boldsymbol{n}, \boldsymbol{g}\rangle_\Gamma, && \forall\, \underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{\Sigma}}_2, && (2.8b)\\
&(\mathbf{div}\underline{\boldsymbol{\sigma}}, \boldsymbol{v}) = -(\boldsymbol{f}, \boldsymbol{v}), && \forall\, \boldsymbol{v} \in \boldsymbol{V}. && (2.8c)
\end{aligned}$$

Introduce the space $\underline{\boldsymbol{\Sigma}} = \underline{\boldsymbol{\Sigma}}_1 \times \underline{\boldsymbol{\Sigma}}_2$ with the norm

$$\|(\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}})\|_{\underline{\boldsymbol{\Sigma}}}^2 = \|\underline{\boldsymbol{s}}\|_0^2 + \|\underline{\boldsymbol{\tau}}\|_{\underline{\boldsymbol{\Sigma}}_2}^2, \quad \forall\, (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}) \in \underline{\boldsymbol{\Sigma}}. \tag{2.9}$$

Then, according to (2.5a) and (2.8), we develop the following augmented formulation: Find $((\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \boldsymbol{u}) \in \underline{\boldsymbol{\Sigma}} \times \boldsymbol{V}$ such that

$$\begin{aligned}
&[\mathcal{A}(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}})] + [\mathcal{B}(\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}), \boldsymbol{u}] = [\mathcal{G}, (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}})], && \forall\, (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}) \in \underline{\boldsymbol{\Sigma}}, && (2.10a)\\
&[\mathcal{B}(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \boldsymbol{v}] = [\mathcal{F}, \boldsymbol{v}], && \forall\, \boldsymbol{v} \in \boldsymbol{V}, && (2.10b)
\end{aligned}$$

where the nonlinear operator $\mathcal{A} : \underline{\boldsymbol{\Sigma}} \to \underline{\boldsymbol{\Sigma}}'$, bilinear operator $\mathcal{B} : \underline{\boldsymbol{\Sigma}} \to \boldsymbol{V}'$, the functionals $\mathcal{F} \in \boldsymbol{V}'$ and $\mathcal{G} \in \underline{\boldsymbol{\Sigma}}'$ are defined by

$$\begin{aligned}
&[\mathcal{A}(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}})] = \left(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}), \underline{\boldsymbol{s}}\right) - (\underline{\boldsymbol{\sigma}}^d, \underline{\boldsymbol{s}}) + (\underline{\boldsymbol{t}}, \underline{\boldsymbol{\tau}}^d)\\
&\qquad\qquad + \kappa\left(\underline{\boldsymbol{\sigma}}^d - 2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}), \underline{\boldsymbol{\tau}}^d\right), && \forall\, (\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}) \in \underline{\boldsymbol{\Sigma}},\\
&[\mathcal{B}(\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}), \boldsymbol{v}] = (\mathbf{div}\underline{\boldsymbol{\tau}}, \boldsymbol{v}), && \forall\, (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}) \in \underline{\boldsymbol{\Sigma}}, \quad \forall\, \boldsymbol{v} \in \boldsymbol{V}, \qquad (2.11)\\
&[\mathcal{F}, \boldsymbol{v}] = -(\boldsymbol{f}, \boldsymbol{v}), && \forall\, \boldsymbol{v} \in \boldsymbol{V},\\
&[\mathcal{G}, (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}})] = \langle \underline{\boldsymbol{\tau}}\boldsymbol{n}, \boldsymbol{g}\rangle_\Gamma, && \forall\, (\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}) \in \underline{\boldsymbol{\Sigma}}.
\end{aligned}$$

Here, the stabilization parameter $\kappa > 0$ is chosen later, $[\cdot,\cdot]$ stands for the duality pairing induced by the operator and the dual space $\underline{\boldsymbol{\Sigma}}'$ is equipped with the following norm:

$$\left\|\mathcal{A}(\underline{\boldsymbol{r}},\underline{\boldsymbol{\xi}})\right\|_{\underline{\boldsymbol{\Sigma}}'} = \sup_{(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\in\underline{\boldsymbol{\Sigma}}\setminus\{\boldsymbol{0}\}} \frac{\left[\mathcal{A}(\underline{\boldsymbol{r}},\underline{\boldsymbol{\xi}}),(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\right]}{\|(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\|_{\underline{\boldsymbol{\Sigma}}}}, \quad \forall\,(\underline{\boldsymbol{r}},\underline{\boldsymbol{\xi}})\in\underline{\boldsymbol{\Sigma}}. \tag{2.12}$$

Other dual norms are defined similarly.

2.3. Well-posedness of the continuous formulation

In this subsection, we show that the problem (2.10) has a unique solution. First, we study the Lipschitz-continuity and monotonicity of $\mathcal{A}$. To do so, we define an auxiliary nonlinear operator $\mathbb{A}:\underline{\boldsymbol{\Sigma}}_1\to\underline{\boldsymbol{\Sigma}}_1'$ given by

$$[\mathbb{A}(\underline{\boldsymbol{\theta}}),\underline{\boldsymbol{s}}] = \int_\Omega \underline{\psi}(\underline{\boldsymbol{\theta}}):\underline{\boldsymbol{s}}\,d\boldsymbol{x}, \quad \forall\,\underline{\boldsymbol{\theta}},\underline{\boldsymbol{s}}\in\underline{\boldsymbol{\Sigma}}_1. \tag{2.13}$$

Then, from (2.2) and (2.3), it shows that the nonlinear operator $\mathbb{A}$ is Lipschitz-continuous and strongly monotone.

Lemma 2.1 ([13, 21, 24, 27]). *Let $r_0 > 0$ and $r_1 > 0$ be the constants in the inequalities* (2.2) *and* (2.3)*, respectively. Then, for any $\underline{\boldsymbol{\theta}}$ and $\underline{\boldsymbol{s}}\in\underline{\boldsymbol{\Sigma}}_1$, it holds*

$$[\mathbb{A}(\underline{\boldsymbol{\theta}})-\mathbb{A}(\underline{\boldsymbol{s}}),\underline{\boldsymbol{\theta}}-\underline{\boldsymbol{s}}] \geq r_0\|\underline{\boldsymbol{\theta}}-\underline{\boldsymbol{s}}\|_0^2, \tag{2.14}$$

$$\|\mathbb{A}(\underline{\boldsymbol{\theta}})-\mathbb{A}(\underline{\boldsymbol{s}})\|_{\underline{\boldsymbol{\Sigma}}_1'} \leq r_1\|\underline{\boldsymbol{\theta}}-\underline{\boldsymbol{s}}\|_0. \tag{2.15}$$

According to Lemma 2.1 and the Cauchy-Schwarz inequality, we can verify the Lipschitz-continuity of $\mathcal{A}$, i.e.

$$\|\mathcal{A}(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}})-\mathcal{A}(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\|_{\underline{\boldsymbol{\Sigma}}'} \leq 3\max\{1,2r_1,\kappa,2\kappa r_1\}\|(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}})-(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\|_{\underline{\boldsymbol{\Sigma}}}, \quad \forall(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}}),(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\in\underline{\boldsymbol{\Sigma}}. \tag{2.16}$$

Then, with the condition of $\kappa\in(0,r_0/r_1^2)$, we obtain the following lemma.

Lemma 2.2. *Let $\mathcal{A}$ and $\mathcal{B}$ be the operators defined by* (2.11)*. And set*

$$\widehat{\underline{\boldsymbol{\Sigma}}} = \mathrm{Ker}(\mathcal{B}) = \underline{\boldsymbol{\Sigma}}_1\times\{\underline{\boldsymbol{\tau}}\in\underline{\boldsymbol{\Sigma}}_2 : \mathbf{div}\underline{\boldsymbol{\tau}}=\boldsymbol{0}\}.$$

Assume that $\kappa\in(0,r_0/r_1^2)$, where r_0 and r_1 are positive constants given in (2.2) *and* (2.3)*. Then,*

$$\begin{aligned}&\left[\mathcal{A}\big((\underline{\boldsymbol{r}},\underline{\boldsymbol{\xi}})+(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}})\big)-\mathcal{A}\big((\underline{\boldsymbol{r}},\underline{\boldsymbol{\xi}})+(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\big),(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}})-(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\right]\\ &\geq \min\left\{2(r_0-\kappa r_1^2),\frac{\kappa}{2}\right\}\|(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}})-(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\|_{\underline{\boldsymbol{\Sigma}}}^2\end{aligned} \tag{2.17}$$

for all $(\underline{\boldsymbol{r}},\underline{\boldsymbol{\xi}})\in\underline{\boldsymbol{\Sigma}}$ and all $(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}}),(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\in\widehat{\underline{\boldsymbol{\Sigma}}}$.

Proof. It is a particular case with symmetric tensors of [24, Lemma 3.2]. □

In the following lemma, we show the inf-sup condition for $[\mathcal{B}(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}}),\boldsymbol{v}]$.

Lemma 2.3. *There exists a constant $\beta > 0$ such that*

$$\sup_{(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\in\underline{\boldsymbol{\Sigma}}\setminus\{\boldsymbol{0}\}} \frac{[\mathcal{B}(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}}),\boldsymbol{v}]}{\|(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\|_{\underline{\boldsymbol{\Sigma}}}} \geq \beta\|\boldsymbol{v}\|_0, \quad \forall\boldsymbol{v}\in\boldsymbol{V}. \tag{2.18}$$

Proof. For any $\boldsymbol{v} \in \boldsymbol{V}$, there exists a $\underline{\boldsymbol{\sigma}}^* \in \underline{\boldsymbol{H}}(\mathbf{div};\mathbb{S})$ and a positive constant $C_0 > 0$ such that [3, 11]

$$\mathbf{div}\underline{\boldsymbol{\sigma}}^* = \boldsymbol{v} \quad \text{in } \Omega, \quad \|\underline{\boldsymbol{\sigma}}^*\|_{\mathbf{div}} \leq C_0\|\boldsymbol{v}\|_0.$$

Set

$$\gamma = \frac{1}{|\Omega|}\int_\Omega \operatorname{tr}(\underline{\boldsymbol{\sigma}}^*)d\boldsymbol{x}, \quad \widetilde{\underline{\boldsymbol{\sigma}}} = \underline{\boldsymbol{\sigma}}^* - \frac{\gamma}{n}\underline{\boldsymbol{I}}.$$

Then, it is straightforward to show that

$$\widetilde{\underline{\boldsymbol{\sigma}}} \in \underline{\boldsymbol{\Sigma}}_2, \quad \mathbf{div}\widetilde{\underline{\boldsymbol{\sigma}}} = \boldsymbol{v} \quad \text{in } \Omega. \tag{2.19}$$

In addition,

$$\begin{aligned}\|\widetilde{\underline{\boldsymbol{\sigma}}}\|_{\mathbf{div}}^2 &= \left(\underline{\boldsymbol{\sigma}}^* - \frac{\gamma}{n}\underline{\boldsymbol{I}}, \underline{\boldsymbol{\sigma}}^* - \frac{\gamma}{n}\underline{\boldsymbol{I}}\right) + \left(\mathbf{div}\left(\underline{\boldsymbol{\sigma}}^* - \frac{\gamma}{n}\underline{\boldsymbol{I}}\right), \mathbf{div}\left(\underline{\boldsymbol{\sigma}}^* - \frac{\gamma}{n}\underline{\boldsymbol{I}}\right)\right) \\ &= (\underline{\boldsymbol{\sigma}}^*, \underline{\boldsymbol{\sigma}}^*) - \frac{1}{n|\Omega|}\left(\int_\Omega \operatorname{tr}(\underline{\boldsymbol{\sigma}}^*)d\boldsymbol{x}\right)^2 + \big(\mathbf{div}(\underline{\boldsymbol{\sigma}}^*), \mathbf{div}(\underline{\boldsymbol{\sigma}}^*)\big) \\ &\leq C_0^2\|\boldsymbol{v}\|_0^2.\end{aligned} \tag{2.20}$$

Note that $\mathcal{B}$ does not depend on $\underline{\boldsymbol{s}}$, so we choose $(\underline{\boldsymbol{s}}, \underline{\boldsymbol{\tau}}) = (\underline{0}, \widetilde{\underline{\boldsymbol{\sigma}}})$ to obtain

$$\sup_{(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\in\underline{\boldsymbol{\Sigma}}\backslash\{\mathbf{0}\}} \frac{[\mathcal{B}(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}}),\boldsymbol{v}]}{\|(\underline{\boldsymbol{s}},\underline{\boldsymbol{\tau}})\|_{\underline{\boldsymbol{\Sigma}}}} \geq \frac{(\mathbf{div}\widetilde{\underline{\boldsymbol{\sigma}}},\boldsymbol{v})}{\|\widetilde{\underline{\boldsymbol{\sigma}}}\|_{\underline{\boldsymbol{\Sigma}}_2}} \geq \frac{\|\boldsymbol{v}\|_0^2}{\|\widetilde{\underline{\boldsymbol{\sigma}}}\|_{\mathbf{div}}} \geq \beta\|\boldsymbol{v}\|_0. \tag{2.21}$$

Here, $\beta = 1/C_0$. The proof is finished. □

To establish the well-posedness of the nonlinear problem (2.10), we recall the following result from nonlinear functional analysis.

Lemma 2.4 ([24, 40]). *Let X, M be Hilbert spaces, $\mathcal{A}: X \to X'$ is a nonlinear operator, and $\mathcal{B}: X \to M'$ is a linear operator. Let $\hat{X} = \operatorname{Ker}(\mathcal{B}) = \{x \in X : [\mathcal{B}(x), q] = 0,\ \forall q \in M\}$. Assume that $\mathcal{A}$ is Lipschitz-continuous on X and that for all $\tilde{z} \in X$, $\mathcal{A}(\tilde{z} + \cdot)$ is uniformly strongly monotone on $\hat{X}$, that is, there exist constants $\rho, \alpha > 0$ such that*

$$\begin{aligned}&\|\mathcal{A}(x) - \mathcal{A}(y)\|_{X'} \leq \rho\|x - y\|_X, && \forall x, y \in X, \\ &[\mathcal{A}(\tilde{z} + x) - \mathcal{A}(\tilde{z} + y), x - y] \geq \alpha\|x - y\|_X^2, && \forall x, y \in \hat{X}, \quad \forall \tilde{z} \in X.\end{aligned}$$

In addition, assume that there exists $\beta > 0$ such that for all $q \in M$,

$$\sup_{x\in X\backslash\{0\}} \frac{[\mathcal{B}(x), q]}{\|x\|_X} \geq \beta\|q\|_M.$$

Then, given $(\mathcal{G}, \mathcal{F}) \in X' \times M'$, there exists a unique $(x, p) \in X \times M$ such that

$$\begin{aligned}&[\mathcal{A}(x), y] + [\mathcal{B}(y), p] = [\mathcal{G}, y], && \forall\, y \in X, \\ &[\mathcal{B}(x), q] = [\mathcal{F}, q], && \forall\, q \in M.\end{aligned}$$

Moreover, the following estimates hold:

$$\begin{aligned}\|x\|_X &\leq \frac{1}{\alpha}\|\mathcal{G}\| + \frac{1}{\beta}\left(1 + \frac{\rho}{\alpha}\right)\|\mathcal{F}\|, \\ \|p\|_M &\leq \frac{1}{\beta}\left(1 + \frac{\rho}{\alpha}\right)\left(\|\mathcal{G}\| + \frac{\rho}{\beta}\|\mathcal{F}\|\right).\end{aligned}$$

According to the Lemma 2.4, we can get the well-posedness of the augmented formulation (2.10) by (2.16), Lemmas 2.2 and 2.3.

Theorem 2.1. *Under the condition of* $\kappa \in (0, r_0/r_1^2)$, *there exists a unique solution* $((\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \boldsymbol{u}) \in \underline{\boldsymbol{\Sigma}} \times \boldsymbol{V}$ *for problem* (2.10).

Remark 2.1. For the case of non-symmetric variables $\underline{\boldsymbol{t}}$ and $\underline{\boldsymbol{\sigma}}$, the well-posedness of the problem, is given by [24, Theorem 3.2].

3. Mixed DG Method

In this section, a stress-strain-velocity-based augmented MDG scheme for the quasi-Newtonian Stokes flows is formulated and its well-posedness is proved.

3.1. DG notation and the MDG scheme

Let $\{\mathcal{T}_h\}_h$ be a family of quasi-regular decomposition of the domain $\overline{\Omega}$ by triangles (tetrahedrons) K. The diameter of K is denoted by h_K, h_e is the length of edge $e \subset \partial K$ and $h = \max\{h_K : K \in \mathcal{T}_h\}$. Denote the union of the boundaries of all the $K \in \mathcal{T}_h$ by $\mathcal{E}_h$. In addition, $\mathcal{E}_h^i$ is the set of all the interior edges and $\mathcal{E}_h^\partial = \mathcal{E}_h/\mathcal{E}_h^i$ is the set of boundary edges. Let $\underline{\boldsymbol{\nabla}}_h$ and $\mathbf{div}_h$ be the broken gradient and divergence operators whose restrictions on each element $K \in \mathcal{T}_h$ are equal to $\underline{\boldsymbol{\nabla}}$ and $\mathbf{div}$, respectively. Next, given an integer $k \geq 0$, we denote by $\mathcal{P}_k(D)$ the space of polynomials defined in D of total degree at most k. Recalling the notation for vector-valued, tensor-valued and symmetric tensor-valued function spaces, we have $\boldsymbol{\mathcal{P}}_k(D) = [\mathcal{P}_k(D)]^n, \underline{\boldsymbol{\mathcal{P}}}_k(D) = [\mathcal{P}_k(D)]^{n\times n}$ and $\underline{\boldsymbol{\mathcal{P}}}_k^{\mathbb{S}}(D) = \{\underline{\boldsymbol{\tau}} \in [\mathcal{P}_k(D)]^{n\times n} : \underline{\boldsymbol{\tau}}^t = \underline{\boldsymbol{\tau}}\}$. Throughout the paper, we use the abbreviation $x \lesssim y$ ($x \gtrsim y$) for the inequality $x \leq Cy$ ($x \geq Cy$), where the letter C denotes a positive constant independent of the mesh size h, but may depend on r_0, r_1 and κ. And it may stand for different values at its different occurrences.

For an interior edge $e \in \mathcal{E}_h^i$ shared by elements K^+ and K^-, let $\boldsymbol{n}^+$ and $\boldsymbol{n}^-$ be the unit normal vectors on e pointing exterior to K^+ and K^-, respectively. We introduce vector-valued functions $\boldsymbol{v}^\pm = \boldsymbol{v}|_{\partial K^\pm}$ and tensor-valued functions $\underline{\boldsymbol{\tau}}^\pm = \underline{\boldsymbol{\tau}}|_{\partial K^\pm}$. Then, we set the averages $\{\cdot\}$ and the jumps $[\![\cdot]\!], [\cdot]$ for $e \in \mathcal{E}_h^i$ as follows:

$$
\begin{aligned}
&\{\boldsymbol{v}\} = \frac{1}{2}(\boldsymbol{v}^+ + \boldsymbol{v}^-), \quad [\![\boldsymbol{v}]\!] = \frac{1}{2}(\boldsymbol{v}^+ \otimes \boldsymbol{n}^+ + \boldsymbol{v}^- \otimes \boldsymbol{n}^- + \boldsymbol{n}^+ \otimes \boldsymbol{v}^+ + \boldsymbol{n}^- \otimes \boldsymbol{v}^-),\\
&\{\underline{\boldsymbol{\tau}}\} = \frac{1}{2}(\underline{\boldsymbol{\tau}}^+ + \underline{\boldsymbol{\tau}}^-), \quad [\underline{\boldsymbol{\tau}}] = \underline{\boldsymbol{\tau}}^+\boldsymbol{n}^+ + \underline{\boldsymbol{\tau}}^-\boldsymbol{n}^-,
\end{aligned}
$$

where $\boldsymbol{v} \otimes \boldsymbol{w}$ is a matrix with $v_i w_j$ as its (i,j)-th element. On boundary edge $e \in \mathcal{E}_h^\partial$, define

$$
\begin{aligned}
&\{\boldsymbol{v}\} = \boldsymbol{v}, \quad [\![\boldsymbol{v}]\!] = \frac{1}{2}(\boldsymbol{v} \otimes \boldsymbol{n} + \boldsymbol{n} \otimes \boldsymbol{v}),\\
&\{\underline{\boldsymbol{\tau}}\} = \underline{\boldsymbol{\tau}}, \quad [\underline{\boldsymbol{\tau}}] = \underline{\boldsymbol{\tau}}\boldsymbol{n}.
\end{aligned}
$$

For any tensor-valued function $\underline{\boldsymbol{\tau}}$ and vector-valued function $\boldsymbol{v}$, a straightforward computation shows that

$$
\sum_{T\in\mathcal{T}_h} \int_{\partial K} \underline{\boldsymbol{\tau}}\boldsymbol{n}_K \cdot \boldsymbol{v} ds = \sum_{e\in\mathcal{E}_h^i} \int_e [\underline{\boldsymbol{\tau}}] \cdot \{\boldsymbol{v}\} ds + \sum_{e\in\mathcal{E}_h} \int_e \{\underline{\boldsymbol{\tau}}\} : [\![\boldsymbol{v}]\!] ds. \tag{3.1}
$$

Denote the discontinuous finite element spaces $\underline{\boldsymbol{\Sigma}}_h^1$, $\underline{\boldsymbol{\Sigma}}_h^2$ and $\boldsymbol{V}_h$ by

$$\underline{\boldsymbol{\Sigma}}_h^1 = \left\{ \underline{\boldsymbol{s}}_h \in \underline{\boldsymbol{L}}^2(\Omega;\mathbb{S}) : \underline{\boldsymbol{s}}_h \in \underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}(K),\ \forall K \in \mathcal{T}_h,\ \operatorname{tr}(\underline{\boldsymbol{s}}_h) = 0 \right\}, \tag{3.2a}$$

$$\underline{\boldsymbol{\Sigma}}_h^2 = \left\{ \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{L}}^2(\Omega;\mathbb{S}) : \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}(K),\ \forall K \in \mathcal{T}_h,\ \int_\Omega \operatorname{tr}(\underline{\boldsymbol{\tau}}_h) d\boldsymbol{x} = 0 \right\}, \tag{3.2b}$$

$$\boldsymbol{V}_h = \left\{ \boldsymbol{v}_h \in \boldsymbol{L}^2(\Omega) : \boldsymbol{v}_h \in \boldsymbol{\mathcal{P}}_k(K),\ \forall K \in \mathcal{T}_h \right\}. \tag{3.2c}$$

Now, let us derive the MDG scheme for problem (2.5). Multiplying (2.5) by test functions $\underline{\boldsymbol{s}}_h$, $\underline{\boldsymbol{\tau}}_h$ and $\boldsymbol{v}_h$, respectively, integrating on any element $K \in \mathcal{T}_h$ and applying the Green's formula, we obtain

$$\begin{aligned}
&\big(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}), \underline{\boldsymbol{s}}_h\big)_K - \big(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\sigma}}^d\big)_K = 0, && \forall \underline{\boldsymbol{s}}_h \in \underline{\boldsymbol{\Sigma}}_h^1, && (3.3a)\\
&\big(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\tau}}_h^d\big)_K + (\mathbf{div}\underline{\boldsymbol{\tau}}_h, \boldsymbol{u})_K - \langle \underline{\boldsymbol{\tau}}_h \boldsymbol{n}_K, \boldsymbol{u}\rangle_{\partial K} = 0, && \forall \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2, && (3.3b)\\
&\big(\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\varepsilon}}_h(\boldsymbol{v}_h)\big)_K - \langle \underline{\boldsymbol{\sigma}} \boldsymbol{n}_K, \boldsymbol{v}_h\rangle_{\partial K} = (\boldsymbol{f}, \boldsymbol{v}_h)_K, && \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. && (3.3c)
\end{aligned}$$

Here, $\underline{\boldsymbol{\varepsilon}}_h(\boldsymbol{v}_h) = (\underline{\boldsymbol{\nabla}}_h \boldsymbol{v}_h + (\underline{\boldsymbol{\nabla}}_h \boldsymbol{v}_h)^t)/2$.

Then, we approximate $\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}$ and $\boldsymbol{u}$ by $\underline{\boldsymbol{t}}_h \in \underline{\boldsymbol{\Sigma}}_h^1$, $\underline{\boldsymbol{\sigma}}_h \in \underline{\boldsymbol{\Sigma}}_h^2$ and $\boldsymbol{u}_h \in \boldsymbol{V}_h$, respectively. Introducing the trace of $\underline{\boldsymbol{\sigma}}$ and $\boldsymbol{u}$ on element edge by the numerical fluxes $\widehat{\underline{\boldsymbol{\sigma}}}_h$ and $\widehat{\boldsymbol{u}}_h$, and summing on all $K \in \mathcal{T}_h$, we get

$$\begin{aligned}
&\big(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}_h), \underline{\boldsymbol{s}}_h\big) - \big(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\sigma}}_h^d\big) = 0, && \forall \underline{\boldsymbol{s}}_h \in \underline{\boldsymbol{\Sigma}}_h^1, && (3.4a)\\
&\big(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\tau}}_h^d\big) + (\mathbf{div}_h\underline{\boldsymbol{\tau}}_h, \boldsymbol{u}_h) - \langle \underline{\boldsymbol{\tau}}_h \boldsymbol{n}_K, \widehat{\boldsymbol{u}}_h\rangle_{\partial \mathcal{T}_h} = 0, && \forall \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2, && (3.4b)\\
&\big(\underline{\boldsymbol{\sigma}}_h, \underline{\boldsymbol{\varepsilon}}_h(\boldsymbol{v}_h)\big) - \langle \widehat{\underline{\boldsymbol{\sigma}}}_h \boldsymbol{n}_K, \boldsymbol{v}_h\rangle_{\partial \mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v}_h), && \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. && (3.4c)
\end{aligned}$$

Applying (3.1) and integrating by parts for (3.4c), we have

$$\begin{aligned}
&\big(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}_h), \underline{\boldsymbol{s}}_h\big) - \big(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\sigma}}_h^d\big) = 0, && \forall \underline{\boldsymbol{s}}_h \in \underline{\boldsymbol{\Sigma}}_h^1, && (3.5a)\\
&\big(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\tau}}_h^d\big) + (\mathbf{div}_h\underline{\boldsymbol{\tau}}_h, \boldsymbol{u}_h) - \big\langle [\underline{\boldsymbol{\tau}}_h], \{\widehat{\boldsymbol{u}}_h\}\big\rangle_{\mathcal{E}_h^i} - \big\langle \{\underline{\boldsymbol{\tau}}_h\}, [\![\widehat{\boldsymbol{u}}_h]\!]\big\rangle_{\mathcal{E}_h} = 0, && \forall \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2, && (3.5b)\\
&-(\mathbf{div}_h\underline{\boldsymbol{\sigma}}_h, \boldsymbol{v}_h) + \big\langle [\underline{\boldsymbol{\sigma}}_h - \widehat{\underline{\boldsymbol{\sigma}}}_h], \{\boldsymbol{v}_h\}\big\rangle_{\mathcal{E}_h^i} + \big\langle \{\underline{\boldsymbol{\sigma}}_h - \widehat{\underline{\boldsymbol{\sigma}}}_h\}, [\![\boldsymbol{v}_h]\!]\big\rangle_{\mathcal{E}_h} = (\boldsymbol{f}, \boldsymbol{v}_h), && \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. && (3.5c)
\end{aligned}$$

We choose the numerical fluxes $\widehat{\underline{\boldsymbol{\sigma}}}_h$ and $\widehat{\boldsymbol{u}}_h$ as

$$\widehat{\underline{\boldsymbol{\sigma}}}_h = \{\underline{\boldsymbol{\sigma}}_h\} \quad \text{and} \quad \widehat{\boldsymbol{u}}_h = \{\boldsymbol{u}_h\} - \frac{\eta_e}{h_e}[\underline{\boldsymbol{\sigma}}_h] \quad \text{on } e \in \mathcal{E}_h^i, \tag{3.6}$$

$$\widehat{\underline{\boldsymbol{\sigma}}}_h = \underline{\boldsymbol{\sigma}}_h \quad \text{and} \quad \widehat{\boldsymbol{u}}_h = \boldsymbol{g} \qquad \text{on } e \in \mathcal{E}_h^\partial, \tag{3.7}$$

where the constant η_e is penalty parameter. With such choices, problem (3.5) can be reformulated as

$$\begin{aligned}
&\big(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}_h), \underline{\boldsymbol{s}}_h\big) - \big(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\sigma}}_h^d\big) = 0, && \forall \underline{\boldsymbol{s}}_h \in \underline{\boldsymbol{\Sigma}}_h^1, && (3.8a)\\
&\big(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\tau}}_h^d\big) + (\mathbf{div}_h\underline{\boldsymbol{\tau}}_h, \boldsymbol{u}_h) - \big\langle [\underline{\boldsymbol{\tau}}_h], \{\boldsymbol{u}_h\}\big\rangle_{\mathcal{E}_h^i} \\
&\quad + \sum_{e\in\mathcal{E}_h^i} \frac{\eta_e}{h_e}\big\langle [\underline{\boldsymbol{\sigma}}_h], [\underline{\boldsymbol{\tau}}_h]\big\rangle_e = \langle \underline{\boldsymbol{\tau}}_h \boldsymbol{n}, \boldsymbol{g}\rangle_{\mathcal{E}_h^\partial}, && \forall \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2, && (3.8b)\\
&(\mathbf{div}_h\underline{\boldsymbol{\sigma}}_h, \boldsymbol{v}_h) - \big\langle [\underline{\boldsymbol{\sigma}}_h], \{\boldsymbol{v}_h\}\big\rangle_{\mathcal{E}_h^i} = -(\boldsymbol{f}, \boldsymbol{v}_h), && \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. && (3.8c)
\end{aligned}$$

In addition, we introduce the space $\underline{\boldsymbol{\Sigma}}_h = \underline{\boldsymbol{\Sigma}}_h^1 \times \underline{\boldsymbol{\Sigma}}_h^2$ and equip it with the following norm:

$$\|(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|^2_{\underline{\boldsymbol{\Sigma}}_h} = \|\underline{\boldsymbol{s}}_h\|_0^2 + \|\underline{\boldsymbol{\tau}}_h\|^2_{\underline{\boldsymbol{\Sigma}}_h^2} = \|\underline{\boldsymbol{s}}_h\|_0^2 + \left\|\underline{\boldsymbol{\tau}}_h^d\right\|_0^2 + \|\mathbf{div}_h \underline{\boldsymbol{\tau}}_h\|_0^2 + |\underline{\boldsymbol{\tau}}_h|_*^2. \tag{3.9}$$

Here,

$$|\underline{\boldsymbol{\tau}}_h|_*^2 = \sum_{e\in\mathcal{E}_h^i} h_e^{-1}\left\|[\underline{\boldsymbol{\tau}}_h]\right\|_{0,e}^2.$$

The following lemma implies that the norm (3.9) is well-defined.

Lemma 3.1 ([26]). *For every $\underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2$, it holds*

$$\|\underline{\boldsymbol{\tau}}_h\|_0^2 \lesssim \left\|\underline{\boldsymbol{\tau}}_h^d\right\|_0^2 + \|\mathbf{div}_h \underline{\boldsymbol{\tau}}_h\|_0^2 + |\underline{\boldsymbol{\tau}}_h|_*^2.$$

For each $(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h$ and each $\boldsymbol{v}_h \in \boldsymbol{V}_h$, define the nonlinear operator $\mathcal{A}_h$ and the linear form $\mathcal{B}_h$ by

$$\begin{aligned}\left[\mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\right] &= \left(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}_h), \underline{\boldsymbol{s}}_h\right) - \left(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\sigma}}_h^d\right) + \left(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\tau}}_h^d\right) \\ &\quad + \sum_{e\in\mathcal{E}_h^i} \frac{\eta_e}{h_e}\left\langle[\underline{\boldsymbol{\sigma}}_h], [\underline{\boldsymbol{\tau}}_h]\right\rangle_e + \kappa\left(\underline{\boldsymbol{\sigma}}_h^d - 2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}_h), \underline{\boldsymbol{\tau}}_h^d\right), \end{aligned} \tag{3.10}$$

$$\left[\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{v}_h\right] = (\mathbf{div}_h \underline{\boldsymbol{\tau}}_h, \boldsymbol{v}_h) - \left\langle[\underline{\boldsymbol{\tau}}_h], \{\boldsymbol{v}_h\}\right\rangle_{\mathcal{E}_h^i}. \tag{3.11}$$

Then, the augmented MDG formulation becomes: Find $((\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), \boldsymbol{u}_h) \in \underline{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h$ such that

$$[\mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)] + [\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{u}_h] = [\mathcal{G}, (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)], \quad \forall\, (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h, \tag{3.12a}$$

$$[\mathcal{B}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), \boldsymbol{v}_h] = [\mathcal{F}, \boldsymbol{v}_h], \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.12b}$$

3.2. Well-posedness of the augmented MDG scheme

In this subsection, we give some inequalities by lemmas and establish the well-posedness of the augmented MDG scheme (3.12).

Firstly, we define the subspace $\widehat{\underline{\boldsymbol{\Sigma}}}_h$ of $\underline{\boldsymbol{\Sigma}}_h$ by

$$\widehat{\underline{\boldsymbol{\Sigma}}}_h = \mathrm{Ker}(\mathcal{B}_h) = \left\{(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h : [\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{v}_h] = 0,\ \forall \boldsymbol{v}_h \in \boldsymbol{V}_h\right\}.$$

To characterize the kernel $\widehat{\underline{\boldsymbol{\Sigma}}}_h$, for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$, we define the lifting operator (cf. [2, 12]) $\boldsymbol{r}_e : \boldsymbol{L}^2(e) \to \boldsymbol{V}_h$ by

$$\int_\Omega \boldsymbol{r}_e(\boldsymbol{\varphi}) \cdot \boldsymbol{v}_h d\mathbf{x} = -\int_e \boldsymbol{\varphi} \cdot \{\boldsymbol{v}_h\} d\boldsymbol{s}, \quad \forall\, e \in \mathcal{E}_h. \tag{3.13}$$

Let $\boldsymbol{r}(\boldsymbol{\varphi}) = \sum_{e\in\mathcal{E}_h^i} \boldsymbol{r}_e(\boldsymbol{\varphi})$. Then,

$$\|\boldsymbol{r}(\boldsymbol{\varphi})\|_0^2 = \left\|\sum_{e\in\mathcal{E}_h^i} \boldsymbol{r}_e(\boldsymbol{\varphi})\right\|_0^2 \le (n+1) \sum_{e\in\mathcal{E}_h^i} \|\boldsymbol{r}_e(\boldsymbol{\varphi})\|_0^2.$$

Lemma 3.2. *For any $(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \widehat{\underline{\boldsymbol{\Sigma}}}_h$, we have $\mathbf{div}_h \underline{\boldsymbol{\tau}}_h + \boldsymbol{r}([\underline{\boldsymbol{\tau}}_h]) = 0$.*

Proof. If $(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \widehat{\underline{\boldsymbol{\Sigma}}}_h$, then for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$,

$$0 = [\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{v}_h] = \int_\Omega \mathbf{div}_h \underline{\boldsymbol{\tau}}_h \cdot \boldsymbol{v}_h d\mathbf{x} - \sum_{e\in\mathcal{E}_h^i} \int_e [\underline{\boldsymbol{\tau}}_h] \cdot \{\boldsymbol{v}_h\} d\boldsymbol{s}$$

$$= \int_\Omega \mathbf{div}_h \underline{\boldsymbol{\tau}}_h \cdot \boldsymbol{v}_h d\mathbf{x} + \int_\Omega \boldsymbol{r}([\underline{\boldsymbol{\tau}}_h]) \cdot \boldsymbol{v}_h d\mathbf{x}$$
$$= \int_\Omega \big(\mathbf{div}_h \underline{\boldsymbol{\tau}}_h + \boldsymbol{r}([\underline{\boldsymbol{\tau}}_h])\big) \cdot \boldsymbol{v}_h d\mathbf{x}.$$

Because $\boldsymbol{v}_h = \mathbf{div}_h \underline{\boldsymbol{\tau}}_h + \boldsymbol{r}([\underline{\boldsymbol{\tau}}_h]) \in \boldsymbol{V}_h$, we conclude that $\mathbf{div}_h \underline{\boldsymbol{\tau}}_h + \boldsymbol{r}([\underline{\boldsymbol{\tau}}_h]) = 0$. $\square$

Lemma 3.3 ([2, 12]). *There exists a positive constant C_1, independent of h, such that*

$$\|\boldsymbol{r}_e(\boldsymbol{\varphi})\|_0 \le C_1 h_e^{-\frac{1}{2}} \|\boldsymbol{\varphi}\|_{0,e}.$$

In order to apply Lemma 2.4 to the augmented MDG formulation (3.12), we need the following lemmas about the continuity and monotonicity of the nonlinear operator $\mathcal{A}_h$.

Lemma 3.4. *For any $(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}} \cup \underline{\boldsymbol{\Sigma}}_h$, the nonlinear operator $\mathcal{A}_h$ satisfies*

$$\big\|\mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - \mathcal{A}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\big\|_{\underline{\boldsymbol{\Sigma}}'_h} \lesssim \big\|(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\big\|_{\underline{\boldsymbol{\Sigma}}_h}.$$

Proof. By the definition of $\mathcal{A}_h$, we obtain

$$\begin{aligned}
&\big[\mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - \mathcal{A}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), (\underline{\boldsymbol{\theta}}_h, \underline{\boldsymbol{\xi}}_h)\big] \\
&= 2\big[\mathbb{A}(\underline{\boldsymbol{t}}_h) - \mathbb{A}(\underline{\boldsymbol{s}}_h), \underline{\boldsymbol{\theta}}_h\big] - \big(\underline{\boldsymbol{\theta}}_h, \underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\tau}}_h^d\big) + \big(\underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\xi}}_h^d\big) \\
&\quad + \sum_{e \in \mathcal{E}_h^i} \frac{\eta_e}{h_e} \big\langle [\underline{\boldsymbol{\sigma}}_h] - [\underline{\boldsymbol{\tau}}_h], [\underline{\boldsymbol{\xi}}_h] \big\rangle_e + \kappa\big(\underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\tau}}_h^d, \underline{\boldsymbol{\xi}}_h^d\big) - 2\kappa\Big[\mathbb{A}(\underline{\boldsymbol{t}}_h) - \mathbb{A}(\underline{\boldsymbol{s}}_h), \underline{\boldsymbol{\xi}}_h^d\Big].
\end{aligned}$$

By using (2.15) and the Cauchy-Schwarz inequality, it follows that

$$\begin{aligned}
&\big[\mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - \mathcal{A}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), (\underline{\boldsymbol{\theta}}_h, \underline{\boldsymbol{\xi}}_h)\big] \\
&\le 2r_1 \|\underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h\|_0 \|\underline{\boldsymbol{\theta}}_h\|_0 + \|\underline{\boldsymbol{\theta}}_h\|_0 \big\|\underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\tau}}_h^d\big\|_0 + \|\underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h\|_0 \big\|\underline{\boldsymbol{\xi}}_h^d\big\|_0 \\
&\quad + \eta |\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h|_* |\underline{\boldsymbol{\xi}}_h|_* + \kappa \big\|\underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\tau}}_h^d\big\|_0 \big\|\underline{\boldsymbol{\xi}}_h^d\big\|_0 + 2r_1\kappa \|\underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h\|_0 \big\|\underline{\boldsymbol{\xi}}_h^d\big\|_0 \\
&\lesssim \big\|(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\big\|_{\underline{\boldsymbol{\Sigma}}_h} \big\|(\underline{\boldsymbol{\theta}}_h, \underline{\boldsymbol{\xi}}_h)\big\|_{\underline{\boldsymbol{\Sigma}}_h},
\end{aligned}$$

where

$$\eta = \max_{e \in \mathcal{E}_h^i} \{\eta_e\}.$$

The proof is complete by the definition (2.12) of the dual norm for $\mathcal{A}_h$. $\square$

Lemma 3.5. *Assume that $\kappa \in (0, r_0/r_1^2)$ and η_0 is a positive constant. Let $\mathcal{A}_h$ be the operator as* (3.10) *defined. For all $(\underline{\boldsymbol{r}}_h, \underline{\boldsymbol{\xi}}_h) \in \underline{\boldsymbol{\Sigma}}_h$ and all $(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \widehat{\underline{\boldsymbol{\Sigma}}}_h$, we have*

$$\begin{aligned}
&\Big[\mathcal{A}_h\big((\underline{\boldsymbol{r}}_h, \underline{\boldsymbol{\xi}}_h) + (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h)\big) - \mathcal{A}_h\big((\underline{\boldsymbol{r}}_h, \underline{\boldsymbol{\xi}}_h) + (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\big), (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\Big] \\
&\gtrsim \|(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}^2, \qquad (3.14)
\end{aligned}$$

$$\big[\mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - \mathcal{A}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\big] \gtrsim \big\|(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\big\|_{\underline{\boldsymbol{\Sigma}}_h}^2. \qquad (3.15)$$

Proof. From the Lemma 3.2, we observe that

$$\widehat{\underline{\boldsymbol{\Sigma}}}_h = \underline{\boldsymbol{\Sigma}}_h^1 \times \big\{\underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2 : \mathbf{div}_h \underline{\boldsymbol{\tau}}_h + \boldsymbol{r}([\underline{\boldsymbol{\tau}}_h]) = 0\big\}. \tag{3.16}$$

Then, given $(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h)$ and $(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \widehat{\underline{\boldsymbol{\Sigma}}}_h$, we obtain

$$\begin{aligned} &\big\|\mathbf{div}_h(\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h)\big\|_0^2 = \big\|\boldsymbol{r}\big([\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h]\big)\big\|_0^2 \\ \le\ & (n+1)\sum_{e\in\mathcal{E}_h^i} \big\|r_e\big([\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h]\big)\big\|_0^2 \le (n+1)C_1^2 |\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h|_*^2. \end{aligned} \tag{3.17}$$

For all $(\underline{\boldsymbol{r}}_h, \underline{\boldsymbol{\xi}}_h) \in \underline{\boldsymbol{\Sigma}}_h$ and all $(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \widehat{\underline{\boldsymbol{\Sigma}}}_h$, we have

$$\big[\mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{t}}_h) - \mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{s}}_h), \underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h\big] = \big[\mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{t}}_h) - \mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{s}}_h), (\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{t}}_h) - (\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{s}}_h)\big].$$

Then, using Lemma 2.1, one finds

$$\begin{aligned} &\Big[\mathcal{A}_h\big((\underline{\boldsymbol{r}}_h, \underline{\boldsymbol{\xi}}_h) + (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h)\big) - \mathcal{A}_h\big((\underline{\boldsymbol{r}}_h, \underline{\boldsymbol{\xi}}_h) + (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\big), (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\Big] \\ =\ & 2\big[\mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{t}}_h) - \mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{s}}_h), (\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{t}}_h) - (\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{s}}_h)\big] \\ &+ \sum_{e\in\mathcal{E}_h^i} \frac{\eta_e}{h_e}\big\langle[\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h], [\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h]\big\rangle_e + \kappa\big\|(\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h)^d\big\|_0^2 \\ &- 2\kappa\big[\mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{t}}_h) - \mathbb{A}(\underline{\boldsymbol{r}}_h + \underline{\boldsymbol{s}}_h), (\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h)^d\big] \\ \ge\ & 2r_0\|\underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h\|_0^2 + \kappa\big\|(\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h)^d\big\|_0^2 + \eta_0|\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h|_*^2 \\ &- \frac{\kappa}{2}\big\|(\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h)^d\big\|_0^2 - 2\kappa r_1^2\|\underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h\|_0^2 \\ \ge\ & 2(r_0 - r_1^2\kappa)\|\underline{\boldsymbol{t}}_h - \underline{\boldsymbol{s}}_h\|_0^2 + \frac{\kappa}{2}\big\|(\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h)^d\big\|_0^2 \\ &+ \frac{\eta_0}{1+(n+1)C_1^2}\|\mathbf{div}(\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h)\|_0^2 + \frac{\eta_0}{1+(n+1)C_1^2}|\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\tau}}_h|_*^2 \\ \ge\ & C\|(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h) - (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}^2. \end{aligned}$$

Here,

$$\eta_0 = \min_{e\in\mathcal{E}_h^i}\{\eta_e\}, \quad C = \min\left\{2(r_0 - r_1^2\kappa), \frac{\kappa}{2}, \frac{\eta_0}{1+(n+1)C_1^2}\right\}.$$

According to the above inequality, it is easy to get (3.15). □

Next, we consider the properties of $\mathcal{B}_h$. By the Cauchy-Schwarz inequality and the trace inequality, we have the boundedness of $\mathcal{B}_h$.

Lemma 3.6 ([41]). *For all $(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}} \cup \underline{\boldsymbol{\Sigma}}_h$, it holds that*

$$\begin{aligned} &[\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{v}_h] \lesssim \|(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}\|\boldsymbol{v}_h\|_0, && \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \\ &[\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{v}] \lesssim \|(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}(\|\boldsymbol{v}\|_0 + h|\boldsymbol{v}|_1), && \forall \boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \cup \boldsymbol{V}_h. \end{aligned}$$

In order to obtain the discrete inf-sup condition, we introduce the space [42]

$$\begin{aligned} \underline{\boldsymbol{\Sigma}}_h^{\mathrm{NC}} = \{&\underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{L}}^2(\Omega;\mathbb{S}) : \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}(K),\ \forall K \in \mathcal{T}_h, \text{ and the moments of} \\ &\underline{\boldsymbol{\tau}}_h\boldsymbol{n} \text{ up to degree } k \text{ are continuous across the interior edges}\}, \\ \mathring{\underline{\boldsymbol{\Sigma}}}_h^{\mathrm{NC}} = &\left\{\underline{\boldsymbol{\tau}}_h : \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^{\mathrm{NC}}, \int_\Omega \mathrm{tr}(\underline{\boldsymbol{\tau}}_h)d\boldsymbol{x} = 0\right\} \subset \underline{\boldsymbol{\Sigma}}_h^2. \end{aligned}$$

For any $\boldsymbol{v}_h \in \boldsymbol{V}_h$, there exists a constant $C_2 > 0$ and $\underline{\boldsymbol{\sigma}}_h^* \in \underline{\boldsymbol{\Sigma}}_h^{\rm NC}$ such that [42]

$$\mathbf{div}_h \underline{\boldsymbol{\sigma}}_h^* = \boldsymbol{v}_h, \quad \|\underline{\boldsymbol{\sigma}}_h^*\|_0^2 + \|\mathbf{div}_h \underline{\boldsymbol{\sigma}}_h^*\|_0^2 + |\underline{\boldsymbol{\sigma}}_h^*|_*^2 \le C_2^2 \|\boldsymbol{v}_h\|_0^2.$$

We are now ready to show the discrete inf-sup condition for $\mathcal{B}_h$.

Lemma 3.7. *For all $\boldsymbol{v}_h \in \boldsymbol{V}_h$, it holds*

$$\sup_{(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h \setminus \{\mathbf{0}\}} \frac{[\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{v}_h]}{\|(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}} \gtrsim \|\boldsymbol{v}_h\|_0. \tag{3.18}$$

Proof. Note that $\mathcal{B}_h$ is independent of $\underline{\boldsymbol{s}}_h$, hence we only need to verity that

$$\sup_{\underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2 \setminus \{\mathbf{0}\}} \frac{(\mathbf{div}_h \underline{\boldsymbol{\tau}}_h, \boldsymbol{v}_h) - \langle [\underline{\boldsymbol{\tau}}_h], \{\boldsymbol{v}_h\} \rangle_{\mathcal{E}_h^i}}{\|\underline{\boldsymbol{\tau}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2}} \gtrsim \|\boldsymbol{v}_h\|_0. \tag{3.19}$$

Let

$$\widetilde{\gamma} = \frac{1}{|\Omega|} \int_\Omega \mathrm{tr}(\underline{\boldsymbol{\sigma}}_h^*) d\boldsymbol{x}, \quad \underline{\widetilde{\boldsymbol{\sigma}}}_h = \underline{\boldsymbol{\sigma}}_h^* - \frac{\widetilde{\gamma}}{n} \underline{\boldsymbol{I}}.$$

Obviously, it shows that

$$\underline{\widetilde{\boldsymbol{\sigma}}}_h \in \mathring{\underline{\boldsymbol{\Sigma}}}_h^{\rm NC}, \quad \mathbf{div}_h \underline{\widetilde{\boldsymbol{\sigma}}}_h = \boldsymbol{v}_h \quad \text{in } \Omega. \tag{3.20}$$

Furthermore, by similar argument for (2.20), we get

$$\|\underline{\widetilde{\boldsymbol{\sigma}}}_h\|_0^2 + \|\mathbf{div}_h \underline{\widetilde{\boldsymbol{\sigma}}}_h\|_0^2 + |\underline{\widetilde{\boldsymbol{\sigma}}}_h|_*^2 \le C_2^2 \|\boldsymbol{v}_h\|_0^2.$$

So, we obtain

$$\sup_{\underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}_h^2 \setminus \{\mathbf{0}\}} \frac{(\mathbf{div}_h \underline{\boldsymbol{\tau}}_h, \boldsymbol{v}_h) - \langle [\underline{\boldsymbol{\tau}}_h], \{\boldsymbol{v}_h\} \rangle_{\mathcal{E}_h^i}}{\|\underline{\boldsymbol{\tau}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2}} \ge \frac{(\mathbf{div}_h \underline{\widetilde{\boldsymbol{\sigma}}}_h, \boldsymbol{v}_h)}{\|\underline{\widetilde{\boldsymbol{\sigma}}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2}} \gtrsim \|\boldsymbol{v}_h\|_0. \tag{3.21}$$

The proof is complete. $\square$

According to the Lemma 2.4, using Lemmas 3.4, 3.5 and 3.7, it shows the well-posedness of the augmented formulation (3.12).

Theorem 3.1. *Assume that $\kappa \in (0, r_0/r_1^2)$ and η_0 is a positive constant. The augmented MDG scheme* (3.12) *has a unique solution $((\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), \boldsymbol{u}_h) \in \underline{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h$.*

4. Error Estimates

In this section, we aim to derive the error estimates for the MDG scheme (3.12).

4.1. Error estimate in energy norm

First, we show the consistency of the MDG scheme.

Lemma 4.1. *Let the solution $((\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \boldsymbol{u}) \in \underline{\boldsymbol{\Sigma}} \times \boldsymbol{H}^1(\Omega)$, then*

$$[\mathcal{A}_h(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}) - \mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)] + [\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{u} - \boldsymbol{u}_h] = 0, \quad \forall (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h, \tag{4.1a}$$

$$[\mathcal{B}_h(\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h), \boldsymbol{v}_h] = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{4.1b}$$

Proof. Since $((\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \boldsymbol{u}) \in \underline{\boldsymbol{\Sigma}} \times \boldsymbol{H}^1(\Omega)$, we have $[\underline{\boldsymbol{\sigma}}] = 0$ and $[\![\boldsymbol{u}]\!] = 0$ on each edge $e \in \mathcal{E}_h^i$. And by the constitutive equation, integrating by parts and (2.5), one finds

$$
\begin{aligned}
&[\mathcal{A}_h(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)] + [\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{u}] \\
&= \big(2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}), \underline{\boldsymbol{s}}_h\big) - \big(\underline{\boldsymbol{\sigma}}^d, \underline{\boldsymbol{s}}_h\big) + \big(\underline{\boldsymbol{\tau}}_h^d, \underline{\boldsymbol{t}}\big) + \kappa\big(\underline{\boldsymbol{\sigma}}^d - 2\underline{\boldsymbol{\psi}}(\underline{\boldsymbol{t}}), \underline{\boldsymbol{\tau}}_h^d\big) + (\mathbf{div}_h \underline{\boldsymbol{\tau}}_h, \boldsymbol{u}) - \big\langle [\underline{\boldsymbol{\tau}}_h], \{\boldsymbol{u}\}\big\rangle_{\mathcal{E}_h^i} \\
&= \big(\underline{\boldsymbol{\tau}}_h^d, \underline{\boldsymbol{t}}\big) - \big(\underline{\boldsymbol{\tau}}_h, \underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})\big) + \big\langle \{\underline{\boldsymbol{\tau}}_h\}, [\![\boldsymbol{u}]\!]\big\rangle_{\mathcal{E}_h} + \big\langle [\underline{\boldsymbol{\tau}}_h], \{\boldsymbol{u}\}\big\rangle_{\mathcal{E}_h^i} - \big\langle [\underline{\boldsymbol{\tau}}_h], \{\boldsymbol{u}\}\big\rangle_{\mathcal{E}_h^i} \\
&= \langle \underline{\boldsymbol{\tau}}_h \boldsymbol{n}, \boldsymbol{g}\rangle_{\mathcal{E}_h^\partial},
\end{aligned}
$$

and

$$
[\mathcal{B}_h(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \boldsymbol{v}_h] = (\mathbf{div}\underline{\boldsymbol{\sigma}}, \boldsymbol{v}_h) = -(\boldsymbol{f}, \boldsymbol{v}_h). \tag{4.2}
$$

Then, we complete the proof by combining (3.12) and (4.2). □

Now, we begin to prove the error estimates in energy norm. Let $\boldsymbol{P}_h^k : \boldsymbol{L}^2(K) \to \boldsymbol{\mathcal{P}}_k(K)$ denote the $\boldsymbol{L}^2$-orthogonal projection defined by

$$
\int_K \big(\boldsymbol{P}_h^k \boldsymbol{v} - \boldsymbol{v}\big) \cdot \boldsymbol{w}_h d\boldsymbol{x} = 0, \quad \forall \boldsymbol{w}_h \in \boldsymbol{\mathcal{P}}_k(K), \quad K \in \mathcal{T}_h. \tag{4.3}
$$

And $\underline{\boldsymbol{L}}^2$-orthogonal projector $\underline{\boldsymbol{P}}_h^k : \underline{\boldsymbol{L}}^2(K) \to \underline{\boldsymbol{\mathcal{P}}}_k(K)$ can be defined similarly. Next, we introduce a projector into the finite element space with a commutativity property.

Lemma 4.2 ([29, Theorem 3.6]). *Set*

$$
\underline{\boldsymbol{D}}_\Pi = \underline{\boldsymbol{H}}(\mathbf{div}; \mathbb{S}) \cap \underline{\boldsymbol{L}}^p(\Omega; \mathbb{S})
$$

with any $p > 2$. There exists a continuous projector $\underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} : \underline{\boldsymbol{D}}_\Pi \to \underline{\boldsymbol{\Sigma}}_h^{\mathrm{NC}}$ such that

$$
\mathbf{div}_h\big(\underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}\big) = \boldsymbol{P}_h^k \mathbf{div}_h \underline{\boldsymbol{\tau}}, \quad \forall \underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{D}}_\Pi.
$$

Moreover, for $0 \le r \le k+1$ and $\underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{H}}^{k+2}(\Omega; \mathbb{S})$, we have

$$
\big\|\underline{\boldsymbol{\tau}} - \underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}\big\|_{0,K} \lesssim h^{r+1} |\underline{\boldsymbol{\tau}}|_{r+1, \widetilde{K}}.
$$

Here, $\widetilde{K}$ is the union of simplices that share a vertex with $K \in \mathcal{T}_h$.

According to Lemma 4.2, for $\underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{D}}_\Pi$, we define a new operator

$$
\widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}} = \underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}} - \frac{\widehat{\gamma}}{n} \underline{\boldsymbol{I}},
$$

where

$$
\widehat{\gamma} = \frac{1}{|\Omega|} \int_\Omega \mathrm{tr}(\underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}) d\boldsymbol{x}.
$$

It is easy to check that

$$
\widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}} \in \mathring{\underline{\boldsymbol{\Sigma}}}_h^{\mathrm{NC}}, \quad \mathbf{div}_h\big(\widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}\big) = \boldsymbol{P}_h^k \mathbf{div}_h \underline{\boldsymbol{\tau}}, \quad \forall \underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{D}}_\Pi. \tag{4.4}
$$

Under the condition that $\int_\Omega \mathrm{tr}(\underline{\boldsymbol{\tau}}) d\boldsymbol{x} = 0$, we can prove that

$$
\begin{aligned}
\big\|\underline{\boldsymbol{\tau}} - \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}\big\|_0 &\le \big\|\underline{\boldsymbol{\tau}} - \underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}\big\|_0 + \frac{1}{(n|\Omega|)^{1/2}} \left| \int_\Omega \mathrm{tr}\big(\underline{\boldsymbol{\tau}} - \underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}\big) d\boldsymbol{x} \right| \\
&\lesssim \big\|\underline{\boldsymbol{\tau}} - \underline{\boldsymbol{\Pi}}_h^{\mathrm{div}} \underline{\boldsymbol{\tau}}\big\|_0 \lesssim h^{r+1} |\underline{\boldsymbol{\tau}}|_{r+1}.
\end{aligned} \tag{4.5}
$$

Theorem 4.1. *Let* $((\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \boldsymbol{u}) \in \underline{\boldsymbol{\Sigma}} \times \boldsymbol{H}^1(\Omega)$ *be the solution of the quasi-Newtonian Stokes flows* (1.1)*, and* $((\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), \boldsymbol{u}_h) \in \underline{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h$ *be the solution of* (3.12)*. Assume that* $\kappa \in (0, r_0/r_1^2), \eta_0$ *is a positive constant and* $(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}, \boldsymbol{u}) \in \underline{\boldsymbol{H}}^{k+1}(\Omega;\mathbb{S}) \times \underline{\boldsymbol{H}}^{k+2}(\Omega;\mathbb{S}) \times \boldsymbol{H}^{k+1}(\Omega)$*, we have*

$$\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0 + \|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_0 \lesssim h^{k+1}\big(|\underline{\boldsymbol{t}}|_{k+1} + |\underline{\boldsymbol{\sigma}}|_{k+2} + |\boldsymbol{u}|_{k+1}\big).$$

Proof. According to (4.1b) and (4.4), we obtain

$$\Big[\mathcal{B}_h\big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\big), \boldsymbol{v}_h\Big] = \big[\mathcal{B}_h(\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h), \boldsymbol{v}_h\big] = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h \tag{4.6}$$

from

$$\Big(\mathbf{div}_h\big(\widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}}\big), \boldsymbol{v}_h\Big) = \big(\boldsymbol{P}_h^k\mathbf{div}_h\underline{\boldsymbol{\sigma}}, \boldsymbol{v}_h\big) = (\mathbf{div}_h\underline{\boldsymbol{\sigma}}, \boldsymbol{v}_h).$$

Taking $\underline{\boldsymbol{s}}_h = (\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h$, $\underline{\boldsymbol{\tau}}_h = \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h$ and $\boldsymbol{v}_h = \boldsymbol{P}_h^k\boldsymbol{u} - \boldsymbol{u}_h$ in (4.1), by (4.6), one finds

$$\begin{aligned}
&\Big[\mathcal{A}_h\big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}}\big) - \mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), \big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\big)\Big] \\
&= \Big[\mathcal{A}_h\big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}}\big) - \mathcal{A}_h(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}), \big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\big)\Big] \\
&\quad + \Big[\mathcal{B}_h\big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\big), \boldsymbol{P}_h^k\boldsymbol{u} - \boldsymbol{u}\Big].
\end{aligned} \tag{4.7}$$

According to (4.7), Lemmas 3.5 and 3.6, we obtain

$$\begin{aligned}
&\big\|\big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}}\big) - (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h)\big\|_{\underline{\boldsymbol{\Sigma}}_h} \\
&\lesssim \big\|(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}) - \big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}}\big)\big\|_{\underline{\boldsymbol{\Sigma}}_h} + \big\|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big\|_0 + h\big|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big|_1.
\end{aligned}$$

From the above inequality, by the triangle inequality, (2.7) and the fact that $\mathrm{tr}(\underline{\boldsymbol{t}}) = 0$, $\mathrm{tr}((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d) = 0$, we deduce that

$$\begin{aligned}
\|(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}) - (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h} &\lesssim \big\|\big((\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}})^d, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}}\big) - (\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}})\big\|_{\underline{\boldsymbol{\Sigma}}_h} + \big\|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big\|_0 + h\big|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big|_1 \\
&\lesssim \big\|\big(\underline{\boldsymbol{P}}_h^k\underline{\boldsymbol{t}}, \widehat{\underline{\boldsymbol{\Pi}}}_h^{\mathrm{div}}\underline{\boldsymbol{\sigma}}\big) - (\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}})\big\|_{\underline{\boldsymbol{\Sigma}}_h} + \big\|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big\|_0 + h\big|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big|_1.
\end{aligned} \tag{4.8}$$

Then, using Lemmas 3.5-3.7 and (4.1a), we arrive at

$$\begin{aligned}
\big\|\boldsymbol{P}_h^k\boldsymbol{u} - \boldsymbol{u}_h\big\|_0 &\lesssim \sup_{(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h \setminus \{\boldsymbol{0}\}} \frac{\big[\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{P}_h^k\boldsymbol{u} - \boldsymbol{u}_h\big]}{\|(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}} \\
&= \sup_{(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h \setminus \{\boldsymbol{0}\}} \frac{\big[\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{P}_h^k\boldsymbol{u} - \boldsymbol{u}\big] + [\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{u} - \boldsymbol{u}_h]}{\|(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}} \\
&= \sup_{(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h) \in \underline{\boldsymbol{\Sigma}}_h \setminus \{\boldsymbol{0}\}} \frac{\big[\mathcal{B}_h(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h), \boldsymbol{P}_h^k\boldsymbol{u} - \boldsymbol{u}\big] - [\mathcal{A}_h(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}) - \mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), (\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)]}{\|(\underline{\boldsymbol{s}}_h, \underline{\boldsymbol{\tau}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h}} \\
&\lesssim \|(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}) - (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h} + \big\|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big\|_0 + h\big|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big|_1.
\end{aligned}$$

Furthermore, the triangle inequality gives

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_0 \lesssim \|(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}) - (\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h)\|_{\underline{\boldsymbol{\Sigma}}_h} + \big\|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big\|_0 + h\big|\boldsymbol{u} - \boldsymbol{P}_h^k\boldsymbol{u}\big|_1. \tag{4.9}$$

Based on (4.5) and the approximation property of $\boldsymbol{P}_h^k$ and $\underline{\boldsymbol{P}}_h^k$, we finish the proof by combining (4.8) and (4.9). □

4.2. Optimal error estimates in $\underline{L}^2$-norm

In this subsection, following the ideas in [41], we show that the $\underline{\boldsymbol{L}}^2$ error estimates for the strain rate $\underline{\boldsymbol{t}}$ and the stress $\underline{\boldsymbol{\sigma}}$, and L^2 error estimate for the pressure p are optimal when the Stokes finite element pair $\boldsymbol{\mathcal{P}}^c_{k+2}$-$\mathcal{P}_{k+1}$ ($k \geq n$) is stable. Here, $\mathcal{P}^c$ represents the conforming polynomial space.

Firstly, based on the classical BDM projection [9] for three-dimension case, on each element $K \in \mathcal{T}_h$, we define a function $\widetilde{\underline{\boldsymbol{\sigma}}}_h \in \underline{\boldsymbol{\mathcal{P}}}_{k+1}(K)$ by $\widehat{\underline{\boldsymbol{\sigma}}}_h$ and $\underline{\boldsymbol{\sigma}}_h$ in (3.4)

$$\int_e (\widetilde{\underline{\boldsymbol{\sigma}}}_h - \widehat{\underline{\boldsymbol{\sigma}}}_h)\boldsymbol{n} \cdot \boldsymbol{v}_h ds = 0, \qquad \forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{P}}_{k+1}(e), \tag{4.10a}$$

$$\int_K (\widetilde{\underline{\boldsymbol{\sigma}}}_h - \underline{\boldsymbol{\sigma}}_h) : \underline{\boldsymbol{\nabla}} \boldsymbol{v}_h d\boldsymbol{x} = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{P}}_k(K), \tag{4.10b}$$

$$\int_K (\widetilde{\underline{\boldsymbol{\sigma}}}_h - \underline{\boldsymbol{\sigma}}_h) : \underline{\boldsymbol{\tau}}_h d\boldsymbol{x} = 0, \qquad \forall \underline{\boldsymbol{\tau}}_h \in \underline{\boldsymbol{\Sigma}}^c_{h,*}(K), \tag{4.10c}$$

where

$$\underline{\boldsymbol{\Sigma}}^c_{h,*}(K) = \left\{ \underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{\mathcal{P}}}_{k+1}(K) : \underline{\boldsymbol{\tau}}\boldsymbol{n} = \boldsymbol{0} \text{ on } e \subset \partial K \text{ and } (\underline{\boldsymbol{\tau}}, \underline{\boldsymbol{\nabla}} \boldsymbol{v}_h)_K = 0, \ \forall\, \boldsymbol{v}_h \in \boldsymbol{\mathcal{P}}_k(K) \right\}.$$

The system (4.10) can be regarded as the row-wise BDM projection [9]. For two-dimension case, we can define similar BDM projection [10]. According to the definition of BDM projection and the fact that the normal component of the numerical trace for the flux is single-valued, we can get the following lemma.

Lemma 4.3 ([38]). *The function $\widetilde{\underline{\boldsymbol{\sigma}}}_h$ in* (4.10) *is well-defined, and*

$$\widetilde{\underline{\boldsymbol{\sigma}}}_h \in \underline{\boldsymbol{\Sigma}}^c_h = \left\{ \underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{H}}(\mathbf{div}) : \underline{\boldsymbol{\tau}}|_K \in \underline{\boldsymbol{\mathcal{P}}}_{k+1}(K)\ \forall K \in \mathcal{T}_h,\ \int_\Omega \mathrm{tr}(\underline{\boldsymbol{\tau}}) d\boldsymbol{x} = 0 \right\}, \tag{4.11}$$

$$\|\underline{\boldsymbol{\sigma}}_h - \widetilde{\underline{\boldsymbol{\sigma}}}_h\|_{0,K} \lesssim h_K^{\frac{1}{2}} \|(\underline{\boldsymbol{\sigma}}_h - \widehat{\underline{\boldsymbol{\sigma}}}_h)\boldsymbol{n}\|_{0,\partial K}. \tag{4.12}$$

Secondly, by the similar argument in [28,41], we symmetrize $\widetilde{\underline{\boldsymbol{\sigma}}}_h$. With the help of the stable Stokes pair $\boldsymbol{\mathcal{P}}^c_{k+2}$-$\mathcal{P}_{k+1}$ ($k \geq n$), one finds the following result. We refer the reader to [41] for detailed discussion.

Lemma 4.4 ([41]). *Assume that the Stokes pair $\boldsymbol{\mathcal{P}}^c_{k+2}$-$\mathcal{P}_{k+1}$ ($k \geq n$) is stable on the decomposition $\mathcal{T}_h$. For $\widetilde{\underline{\boldsymbol{\sigma}}}_h$ given in* (4.10)*, there exists $\widetilde{\underline{\boldsymbol{\tau}}}_h \in \underline{\boldsymbol{\Sigma}}^c_h$ such that $\underline{\boldsymbol{\sigma}}_{h,*} = \widetilde{\underline{\boldsymbol{\sigma}}}_h + \widetilde{\underline{\boldsymbol{\tau}}}_h \in \underline{\boldsymbol{H}}(\mathbf{div}; \mathbb{S})$,*

$$\mathbf{div} \widetilde{\underline{\boldsymbol{\tau}}}_h = 0, \quad \|\widetilde{\underline{\boldsymbol{\tau}}}_h\|_0 \lesssim \|\underline{\boldsymbol{\sigma}}_h - \widetilde{\underline{\boldsymbol{\sigma}}}_h\|_0. \tag{4.13}$$

Thirdly, the conforming mixed element $\underline{\boldsymbol{\mathcal{P}}}^c_{k+1}$-$\boldsymbol{\mathcal{P}}_k$ ($k \geq n$) on simplicial grids is constructed in [33,34]. Moreover, when $k \geq n$, there exists a projection $\underline{\boldsymbol{\Pi}}^c_h$ [32] such that $\underline{\boldsymbol{\Pi}}^c_h \underline{\boldsymbol{\tau}}$ is symmetric and

$$\left(\mathbf{div}(\underline{\boldsymbol{\tau}} - \underline{\boldsymbol{\Pi}}^c_h \underline{\boldsymbol{\tau}}), \boldsymbol{v}_h\right) = 0, \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.14a}$$

$$\left\|\underline{\boldsymbol{\tau}} - \underline{\boldsymbol{\Pi}}^c_h \underline{\boldsymbol{\tau}}\right\|_0 \lesssim h^{k+2} |\underline{\boldsymbol{\tau}}|_{k+2}, \quad \forall \underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{H}}^{k+2}(\Omega; \mathbb{S}). \tag{4.14b}$$

We refer to [32–34] for detail discussions.

Now, we begin to show the optimal convergence order for $\underline{\boldsymbol{t}}$ and $\underline{\boldsymbol{\sigma}}^d$ in $\underline{\boldsymbol{L}}^2$-norm.

Theorem 4.2. *Assume that the parameter $\kappa \in (0, r_0/r_1^2)$, where r_0 and r_1 are positive constants given in* (2.2) *and* (2.3). *Let $(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}, \boldsymbol{u}) \in \underline{\boldsymbol{H}}^{k+2}(\Omega;\mathbb{S}) \times \underline{\boldsymbol{H}}^{k+2}(\Omega;\mathbb{S}) \times \boldsymbol{H}^{k+1}(\Omega)$ and $((\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), \boldsymbol{u}_h)$ be the solutions of the problem* (2.10) *and the MDG scheme* (3.12), *respectively, we have*

$$\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0 + \left\|\underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d\right\|_0 \lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+2} + |\underline{\boldsymbol{\sigma}}|_{k+2} + |\boldsymbol{u}|_{k+1}). \tag{4.15}$$

Proof. Applying (3.4c), (4.10) and Lemma 4.3, for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$, we have

$$\begin{aligned}(\boldsymbol{f}, \boldsymbol{v}_h) &= \big(\underline{\boldsymbol{\sigma}}_h, \underline{\boldsymbol{\varepsilon}}_h(\boldsymbol{v}_h)\big) - \langle \widehat{\underline{\boldsymbol{\sigma}}}_h \boldsymbol{n}_K, \boldsymbol{v}_h\rangle_{\partial\mathcal{T}_h} \\ &= (\underline{\boldsymbol{\sigma}}_h, \underline{\boldsymbol{\nabla}}_h \boldsymbol{v}_h) - \langle \widehat{\underline{\boldsymbol{\sigma}}}_h \boldsymbol{n}_K, \boldsymbol{v}_h\rangle_{\partial\mathcal{T}_h} \\ &= (\widetilde{\underline{\boldsymbol{\sigma}}}_h, \underline{\boldsymbol{\nabla}}_h \boldsymbol{v}_h) - \langle \widehat{\underline{\boldsymbol{\sigma}}}_h \boldsymbol{n}_K, \boldsymbol{v}_h\rangle_{\partial\mathcal{T}_h} \\ &= -(\mathbf{div}\widetilde{\underline{\boldsymbol{\sigma}}}_h, \boldsymbol{v}_h).\end{aligned} \tag{4.16}$$

By Lemma 4.4, there exist $\widetilde{\underline{\boldsymbol{\tau}}}_h \in \underline{\boldsymbol{\Sigma}}_h^c$ such that the symmetrized variable $\underline{\boldsymbol{\sigma}}_{h,*} = \widetilde{\underline{\boldsymbol{\sigma}}}_h + \widetilde{\underline{\boldsymbol{\tau}}}_h$ is piecewise $\underline{\boldsymbol{\mathcal{P}}}_{k+1}(K)$ and $\underline{\boldsymbol{\sigma}}_{h,*} \in \underline{\boldsymbol{H}}(\mathbf{div};\mathbb{S})$. Then,

$$(\mathbf{div}\underline{\boldsymbol{\sigma}}_{h,*}, \boldsymbol{v}_h) = -(\boldsymbol{f}, \boldsymbol{v}_h). \tag{4.17}$$

From (4.14), we introduce a new projection

$$\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} = \underline{\boldsymbol{\Pi}}_h^c \underline{\boldsymbol{\sigma}} - \frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(\underline{\boldsymbol{\Pi}}_h^c \underline{\boldsymbol{\sigma}}) d\boldsymbol{x} \underline{\boldsymbol{I}}$$

for any $\underline{\boldsymbol{\sigma}} \in \underline{\boldsymbol{H}}^1(\Omega;\mathbb{S})$. It is easy to check that

$$\Big(\mathbf{div}\big(\underline{\boldsymbol{\sigma}} - \widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}}\big), \boldsymbol{v}_h\Big) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.18a}$$

$$\int_\Omega \mathrm{tr}\big(\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}}\big) d\boldsymbol{x} = 0. \tag{4.18b}$$

In addition, noting that $\int_\Omega \mathrm{tr}(\underline{\boldsymbol{\sigma}}) d\boldsymbol{x} = 0$ and by similar argument in (4.5), we get

$$\big\|\underline{\boldsymbol{\sigma}} - \widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}}\big\|_0 \lesssim h^{k+2}|\underline{\boldsymbol{\sigma}}|_{k+2}, \quad \forall \underline{\boldsymbol{\sigma}} \in \underline{\boldsymbol{H}}^{k+2}(\Omega;\mathbb{S}). \tag{4.19}$$

By (4.2), (4.17) and (4.18a), we have

$$\Big(\mathbf{div}\big(\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*}\big), \boldsymbol{v}_h\Big) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{4.20}$$

And it is easy to verify that $\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*} \in \underline{\boldsymbol{\Sigma}}_h^2 \cap \underline{\boldsymbol{H}}(\mathbf{div})$. According to $\mathrm{tr}((\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d) = 0$, we obtain $(\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h \in \underline{\boldsymbol{\Sigma}}_h^1$. Taking $\underline{\boldsymbol{s}}_h = (\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h$ and $\underline{\boldsymbol{\tau}}_h = \widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*}$ in (4.1a), by the $\underline{\boldsymbol{H}}(\mathbf{div})$ conformity of $\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*}$ and the $\boldsymbol{L}^2$-orthogonal projection (4.3) for $\boldsymbol{u}$, it holds

$$\Big[\mathcal{A}_h(\underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}) - \mathcal{A}_h(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h), \big((\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}_h, \widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*}\big)\Big] = 0, \tag{4.21}$$

which implies that

$$\begin{aligned}&2r_0\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0^2 + \kappa\big\|\underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d\big\|_0^2 \\ \le\ & \big|2\big[\mathbb{A}(\underline{\boldsymbol{t}}) - \mathbb{A}(\underline{\boldsymbol{t}}_h), (\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}\big]\big| + \big|\big((\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d - \underline{\boldsymbol{t}}, \underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d\big)\big| \\ &+ \big|\big(\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h, (\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})^d\big)\big| + \big|\big(\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\sigma}}_{h,*}^d\big)\big| + \big|\kappa\big(\underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d, (\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})^d\big)\big| \\ &+ \big|\kappa\big(\underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d, \underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\sigma}}_{h,*}^d\big)\big| + \big|2\kappa\big[\mathbb{A}(\underline{\boldsymbol{t}}) - \mathbb{A}(\underline{\boldsymbol{t}}_h), (\widehat{\underline{\boldsymbol{\Pi}}}_h^c \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})^d\big]\big| \\ &+ \big|2\kappa\big[\mathbb{A}(\underline{\boldsymbol{t}}) - \mathbb{A}(\underline{\boldsymbol{t}}_h), \underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d\big]\big| + \big|2\kappa\big[\mathbb{A}(\underline{\boldsymbol{t}}) - \mathbb{A}(\underline{\boldsymbol{t}}_h), \underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\sigma}}_{h,*}^d\big]\big|.\end{aligned} \tag{4.22}$$

According to Lemma 2.1 and the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
&\left|2\left[\mathbb{A}(\underline{\boldsymbol{t}})-\mathbb{A}(\underline{\boldsymbol{t}}_h),(\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d-\underline{\boldsymbol{t}}\right]\right| \le \frac{r_0}{10}\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0^2+\frac{10r_1^2}{r_0}\left\|(\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d-\underline{\boldsymbol{t}}\right\|_0^2,\\
&\left|\left((\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d-\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\right)\right| \le \frac{\kappa}{18}\left\|\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\right\|_0^2+\frac{9}{2\kappa}\left\|(\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d-\underline{\boldsymbol{t}}\right\|_0^2,\\
&\left|\left(\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h,(\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}})^d\right)\right| \le \frac{r_0}{10}\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0^2+\frac{5}{2r_0}\left\|(\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}})^d\right\|_0^2,\\
&\left|\left(\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h,\underline{\boldsymbol{\sigma}}_h^d-\underline{\boldsymbol{\sigma}}_{h,*}^d\right)\right| \le \frac{r_0}{10}\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0^2+\frac{5}{2r_0}\left\|\underline{\boldsymbol{\sigma}}_h^d-\underline{\boldsymbol{\sigma}}_{h,*}^d\right\|_0^2,\\
&\left|\kappa\left(\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d,(\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}})^d\right)\right| \le \frac{\kappa}{18}\left\|\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\right\|_0^2+\frac{9\kappa}{2}\left\|(\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}})^d\right\|_0^2,\\
&\left|\kappa\left(\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d,\underline{\boldsymbol{\sigma}}_h^d-\underline{\boldsymbol{\sigma}}_{h,*}^d\right)\right| \le \frac{\kappa}{18}\left\|\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\right\|_0^2+\frac{9\kappa}{2}\left\|\underline{\boldsymbol{\sigma}}_h^d-\underline{\boldsymbol{\sigma}}_{h,*}^d\right\|_0^2,\\
&\left|2\kappa\left[\mathbb{A}(\underline{\boldsymbol{t}})-\mathbb{A}(\underline{\boldsymbol{t}}_h),(\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}})^d\right]\right| \le \frac{r_0}{10}\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0^2+\frac{10r_1^2\kappa^2}{r_0}\left\|(\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}})^d\right\|_0^2,\\
&\left|2\kappa\left[\mathbb{A}(\underline{\boldsymbol{t}})-\mathbb{A}(\underline{\boldsymbol{t}}_h),\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\right]\right| \le \frac{2\kappa}{3}\left\|\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\right\|_0^2+\frac{3\kappa r_1^2}{2}\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0^2,\\
&\left|2\kappa\left[\mathbb{A}(\underline{\boldsymbol{t}})-\mathbb{A}(\underline{\boldsymbol{t}}_h),\underline{\boldsymbol{\sigma}}_h^d-\underline{\boldsymbol{\sigma}}_{h,*}^d\right]\right| \le \frac{r_0}{10}\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0^2+\frac{10r_1^2\kappa^2}{r_0}\left\|\underline{\boldsymbol{\sigma}}_h^d-\underline{\boldsymbol{\sigma}}_{h,*}^d\right\|_0^2.
\end{aligned}
$$

Due to the fact that $\mathrm{tr}((\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d)=0$, combining the above inequalities with (4.22) and using (2.7), it holds

$$
\begin{aligned}
&\left(\frac{3r_0}{2}-\frac{3\kappa r_1^2}{2}\right)\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0^2+\frac{\kappa}{6}\left\|\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\right\|_0^2\\
\lesssim\ &\left\|(\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}})^d-\underline{\boldsymbol{t}}\right\|_0^2+\left\|(\underline{\boldsymbol{\sigma}}-\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}})^d\right\|_0^2+\left\|(\underline{\boldsymbol{\sigma}}_h-\underline{\boldsymbol{\sigma}}_{h,*})^d\right\|_0^2\\
\lesssim\ &\left\|\underline{\boldsymbol{P}}_h^{k+1}\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}\right\|_0^2+\left\|\underline{\boldsymbol{\sigma}}-\widehat{\boldsymbol{\Pi}}_h^c\underline{\boldsymbol{\sigma}}\right\|_0^2+\|\underline{\boldsymbol{\sigma}}_h-\underline{\boldsymbol{\sigma}}_{h,*}\|_0^2.
\end{aligned}
$$

In addition, by Theorem 4.1, Lemmas 4.3 and 4.4, we have

$$
\begin{aligned}
\|\underline{\boldsymbol{\sigma}}_h-\underline{\boldsymbol{\sigma}}_{h,*}\|_0 &\le \|\underline{\boldsymbol{\sigma}}_h-\underline{\widetilde{\boldsymbol{\sigma}}}_h\|_0+\|\underline{\widetilde{\boldsymbol{\tau}}}_h\|_0\\
&\lesssim \sum_{e\in\mathcal{E}_h^i} h^{\frac{1}{2}}\|(\underline{\widehat{\boldsymbol{\sigma}}}_h-\underline{\boldsymbol{\sigma}}_h)\boldsymbol{n}\|_{0,e}\lesssim h|\underline{\boldsymbol{\sigma}}_h|_*\\
&\lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+1}+|\underline{\boldsymbol{\sigma}}|_{k+2}+|\boldsymbol{u}|_{k+1}),
\end{aligned}
\tag{4.23}
$$

then choosing $\kappa\in(0,r_0/r_1^2)$ and applying (4.19), we get

$$
\|\underline{\boldsymbol{t}}-\underline{\boldsymbol{t}}_h\|_0+\|\underline{\boldsymbol{\sigma}}^d-\underline{\boldsymbol{\sigma}}_h^d\|_0\lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+2}+|\underline{\boldsymbol{\sigma}}|_{k+2}+|\boldsymbol{u}|_{k+1}),\tag{4.24}
$$

which finishes the proof. □

Finally, the optimal error estimates for $\underline{\boldsymbol{\sigma}}$ in $\underline{\boldsymbol{L}}^2$-norm and p in L^2-norm are given.

Theorem 4.3. *Assume that the solution of* (1.1) *satisfies*

$$
(\underline{\boldsymbol{t}},\underline{\boldsymbol{\sigma}},\boldsymbol{u})\in\underline{\boldsymbol{H}}^{k+2}(\Omega;\mathbb{S})\times\underline{\boldsymbol{H}}^{k+2}(\Omega;\mathbb{S})\times\boldsymbol{H}^{k+1}(\Omega).
$$

Under the condition of Theorem 4.2, the solution of the MDG problem (3.12) *satisfies*

$$
\|\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}}_h\|_0\lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+2}+|\underline{\boldsymbol{\sigma}}|_{k+2}+|\boldsymbol{u}|_{k+1}).\tag{4.25}
$$

In addition, we have the optimal error estimate for pressure p in L^2-norm

$$\|p - p_h\|_0 \lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+2} + |\underline{\boldsymbol{\sigma}}|_{k+2} + |\boldsymbol{u}|_{k+1}). \tag{4.26}$$

Proof. By (4.19), (4.23) and (4.24), we have

$$\begin{aligned}
\left\|(\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}})^d - \underline{\boldsymbol{\sigma}}_{h,*}^d\right\|_0 &\le \left\|(\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}})^d - \underline{\boldsymbol{\sigma}}^d\right\|_0 + \left\|\underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d\right\|_0 + \left\|\underline{\boldsymbol{\sigma}}_h^d - \underline{\boldsymbol{\sigma}}_{h,*}^d\right\|_0 \\
&\le \left\|\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\right\|_0 + \left\|\underline{\boldsymbol{\sigma}}^d - \underline{\boldsymbol{\sigma}}_h^d\right\|_0 + \|\underline{\boldsymbol{\sigma}}_h - \underline{\boldsymbol{\sigma}}_{h,*}\|_0 \\
&\lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+2} + |\underline{\boldsymbol{\sigma}}|_{k+2} + |\boldsymbol{u}|_{k+1}).
\end{aligned}$$

Noting that $\mathbf{div}_h(\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*}) \in \boldsymbol{V}_h$ and taking $\boldsymbol{v}_h = \mathbf{div}_h(\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*})$ in (4.20), we obtain

$$\left\|\mathbf{div}_h(\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*})\right\|_0 = 0. \tag{4.27}$$

Using Lemma 3.1 and (4.27), one gets

$$\begin{aligned}
\left\|\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*}\right\|_0 &\lesssim \left\|(\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*})^d\right\|_0 + \left\|\mathbf{div}_h(\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*})\right\|_0 \\
&\lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+2} + |\underline{\boldsymbol{\sigma}}|_{k+2} + |\boldsymbol{u}|_{k+1}),
\end{aligned}$$

from (4.23), it also yields

$$\begin{aligned}
\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0 &\le \left\|\underline{\boldsymbol{\sigma}} - \widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}}\right\|_0 + \left\|\widehat{\underline{\boldsymbol{\Pi}}}_h^c\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_{h,*}\right\|_0 + \|\underline{\boldsymbol{\sigma}}_{h,*} - \boldsymbol{\sigma}_h\|_0 \\
&\lesssim h^{k+2}(|\underline{\boldsymbol{t}}|_{k+2} + |\underline{\boldsymbol{\sigma}}|_{k+2} + |\boldsymbol{u}|_{k+1}).
\end{aligned}$$

Noticing that (2.4), we can define the numerical solution of pressure by the postprocessed approximation $p_h = -\mathrm{tr}(\underline{\boldsymbol{\sigma}}_h)/n$. The the optimal L^2 estimate for pressure follows from the fact that

$$\|p - p_h\|_0 \lesssim \|\mathrm{tr}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h)\|_0 \lesssim \|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0,$$

which gets (4.26). □

Remark 4.1. When the viscosity μ in (1.1) is a positive constant, assuming that the parameter $\kappa \in (0, 1/(2\mu))$ and $\eta_0 > 0$, we can also obtain the well-posedness the discrete augmented formulation (3.12). And convergence order for the stress in broken $\underline{\boldsymbol{H}}(\mathbf{div})$-norm and velocity in $\boldsymbol{L}^2$-norm are optimal, which are shown in Theorem 4.1. In addition, Theorems 4.2 and 4.3 hold, i.e., the strain rate and the stress in $\underline{\boldsymbol{L}}^2$-norm, and the pressure in L^2-norm are optimal under certain conditions.

5. Numerical Examples

In this section, some numerical examples are implemented by using Fenics [36] to illustrate the performance of the MDG method (3.12). The Newton iteration solver is employed for the nonlinear discrete algebraic system. For simplicity, the uniform triangular meshes in the two-dimensional case and the DG element triplets $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$- $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$, $k = 0, 1, 2$ are considered for all the numerical examples.

For the finite element triplet $\underline{\boldsymbol{\mathcal{P}}}_1^{\mathbb{S}}$- $\underline{\boldsymbol{\mathcal{P}}}_1^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_0$, Example 5.1 is utilized to show the behavior of MDG scheme (3.12) with different nonlinear parameters μ. Examples 5.2 and 5.3 are employed to verify the performance of the MDG scheme with the finite element triplets $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$- $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$, $k = 0, 1, 2$ for linear Stokes and quasi-Newtonian Stokes flows, respectively.

Example 5.1. Let the domain $\Omega = (0,1) \times (0,1)$. The right-hand side function $\boldsymbol{f}$, the exact strain rate function $\underline{\boldsymbol{t}}$, the exact stress function $\underline{\boldsymbol{\sigma}}$ and boundary condition $\boldsymbol{g}$ follow from the exact solution

$$\begin{cases} u_1(x,y) = -\cos(\pi x)\sin(\pi y), \\ u_2(x,y) = \sin(\pi x)\cos(\pi y), \\ p(x,y) = \sin(\pi x) - \dfrac{2}{\pi}. \end{cases}$$

This example aims at testing the accuracy and reliability of the MDG method for the nonlinear kinematic viscosity. The Carreau's law

$$\mu(\xi) = \mu_0 + \mu_1(1+\xi^2)^{\frac{\beta-2}{2}}, \quad \forall \xi \in \mathbb{R}^+$$

with $\mu_0 > 0, \mu_1 > 0$ and $1 \le \beta \le 2$ is adopted and different parameters are given as follows:

$$\begin{aligned} &\mu_0 = 1.0, \quad \mu_1 = 0.6, \quad \beta = 1.0, \\ &\mu_0 = 1.0, \quad \mu_1 = 0.6, \quad \beta = 1.5, \\ &\mu_0 = 1.0, \quad \mu_1 = 0.6, \quad \beta = 2.0. \end{aligned} \tag{5.1}$$

When $\beta = 2.0$, it recovers the usual linear Stokes model. While, it is easy to check that the assumptions (2.2) and (2.3) are satisfied with

$$r_0 = \mu_0, \quad r_1 = \mu_0 + \mu_1\left(\frac{|\beta-2|}{2} + 1\right).$$

We consider the parameter $\kappa = 0.1$, which meets the need $\kappa \in (0, r_0/r_1^2)$ in Theorems 4.1 and 4.2.

The MDG finite element triplet $\underline{\boldsymbol{\mathcal{P}}}_1^{\mathbb{S}}$- $\underline{\boldsymbol{\mathcal{P}}}_1^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_0$ is employed in the numerical discretization on uniform meshes with $1/h = 4, 8, 16, 32$. Then, for different group of parameters listed in (5.1), the numerical results of $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2}$, $\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$, $\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0$, $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0$ and $\|p - p_h\|_0$ are shown in Table 5.1. We can see that all the convergence orders are linear, which confirms the theoretical results.

Table 5.1: Numerical errors and orders in Example 5.1 for $\underline{\boldsymbol{\mathcal{P}}}_1^{\mathbb{S}}$ -$\underline{\boldsymbol{\mathcal{P}}}_1^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_0$ element.

β	$1/h$	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2}$	Order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$	Order	$\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0$	Order	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0$	Order	$\|p - p_h\|_0$	Order
1.0	4	4.3148e+00	—	1.8188e-01	—	1.5364e-01	—	4.7803e-01	—	2.2007e-01	—
	8	2.4962e+00	0.79	9.2257e-02	0.98	6.5984e-02	1.22	2.3272e-01	1.04	1.1348e-01	0.96
	16	1.3031e+00	0.94	4.6294e-02	0.99	3.1489e-02	1.07	1.1512e-01	1.02	5.7181e-02	0.99
	32	6.5715e-01	0.99	2.3168e-02	1.00	1.5596e-02	1.01	5.7555e-02	1.00	2.8730e-02	0.99
1.5	4	4.9247e+00	—	1.8187e-01	—	1.5200e-01	—	5.3189e-01	—	2.4729e-01	—
	8	2.7088e+00	0.86	9.2253e-02	0.98	6.5528e-02	1.21	2.5339e-01	1.07	1.2408e-01	0.99
	16	1.3886e+00	0.96	4.6292e-02	0.99	3.1510e-02	1.06	1.2545e-01	1.01	6.2399e-02	0.99
	32	6.9785e-01	0.99	2.3167e-02	1.00	1.5625e-02	1.01	6.2627e-02	1.00	3.1275e-02	1.00
2.0	4	6.3640e+00	—	1.8189e-01	—	1.5336e-01	—	6.6679e-01	—	3.1920e-01	—
	8	3.2711e+00	0.96	9.2256e-02	0.98	6.7077e-02	1.19	3.0229e-01	1.14	1.5051e-01	1.08
	16	1.6289e+00	1.01	4.6292e-02	0.99	3.2411e-02	1.05	1.4667e-01	1.04	7.3336e-02	1.04
	32	8.1280e-01	1.00	2.3166e-02	1.00	1.6079e-02	1.01	7.2771e-02	1.01	3.6388e-02	1.01

Example 5.2. Set the domain $\Omega = (0,1)\times(0,1)$. The right-hand side function $\boldsymbol{f}$, the exact strain rate function $\underline{\boldsymbol{t}}$, the exact stress function $\underline{\boldsymbol{\sigma}}$ and boundary condition $\boldsymbol{g}$ are selected such that the exact solution is given by

$$\begin{cases} u_1(x,y) = x^2(x-1)^2y(y-1)(2y-1), \\ u_2(x,y) = -x(x-1)(2x-1)y^2(y-1)^2, \\ p(x,y) = (2x-1)(2y-1). \end{cases}$$

In this example, we test the accuracy and reliability of the MDG method for the linear Stokes flow with $\mu = 1,\ 0.1, 0.01$, respectively. Set $1/h = 4, 8, 16, 32$, and $\kappa = 0.1$. We compute the numerical solutions $(\underline{\boldsymbol{t}}_h, \underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h)$ on uniform meshes with the finite element triplets $\underline{\boldsymbol{\mathcal{P}}}^{\mathbb{S}}_{k+1}$-$\underline{\boldsymbol{\mathcal{P}}}^{\mathbb{S}}_{k+1}$-$\boldsymbol{\mathcal{P}}_k$, $k = 0,1,2$. For $\mu = 1, 0.1, 0.01$, the numerical results of $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}^2_h}$, $\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$, $\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0$, $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0$ and $\|p - p_h\|_0$ are given in Tables 5.2, 5.3 and 5.4, respectively. The numerical results are consistent with Theorems 4.2-4.3 and the comments in Remark 4.1.

Table 5.2: Numerical errors and orders in Example 5.2 for $\underline{\boldsymbol{\mathcal{P}}}^{\mathbb{S}}_{k+1}$-$\underline{\boldsymbol{\mathcal{P}}}^{\mathbb{S}}_{k+1}$-$\boldsymbol{\mathcal{P}}_k$ element and $\mu = 1$.

k	$1/h$	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}^2_h}$	Order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$	Order	$\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0$	Order	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0$	Order	$\|p - p_h\|_0$	Order
0	4	4.7934e-01	—	3.2001e-03	—	2.2745e-02	—	7.2854e-02	—	4.0239e-02	—
	8	2.4649e-01	0.96	1.6670e-03	0.94	1.1643e-02	0.97	3.4166e-02	1.09	1.7678e-02	1.19
	16	1.2423e-01	0.99	8.4006e-04	0.99	5.8247e-03	1.00	1.6642e-02	1.04	8.4034e-03	1.07
	32	6.2248e-02	1.00	4.2084e-04	1.00	2.9109e-03	1.00	8.2548e-03	1.01	4.1381e-03	1.02
1	4	3.8937e-02	—	4.3667e-04	—	7.5401e-04	—	2.2053e-03	—	1.1378e-03	—
	8	1.1233e-02	1.79	1.1575e-04	1.92	1.4616e-04	2.37	3.8753e-04	2.51	1.7990e-04	2.66
	16	2.9970e-03	1.91	2.9396e-05	1.98	2.6635e-05	2.46	6.9551e-05	2.48	3.1620e-05	2.51
	32	7.6844e-04	1.96	7.3789e-06	1.99	5.4006e-06	2.30	1.4622e-05	2.25	6.9691e-06	2.18
2	4	8.6430e-03	—	7.5682e-05	—	1.0167e-04	—	2.7561e-04	—	1.3155e-04	—
	8	1.1537e-03	2.91	1.0069e-05	2.91	8.3204e-06	3.61	2.2525e-05	3.61	1.0734e-05	3.62
	16	1.4474e-04	2.99	1.2816e-06	2.97	6.0143e-07	3.79	1.6224e-06	3.80	7.6979e-07	3.80
	32	1.7993e-05	3.01	1.6095e-07	2.99	4.0299e-08	3.90	1.0850e-07	3.90	5.1362e-08	3.91

Table 5.3: Numerical errors and orders in Example 5.2 for $\underline{\boldsymbol{\mathcal{P}}}^{\mathbb{S}}_{k+1}$-$\underline{\boldsymbol{\mathcal{P}}}^{\mathbb{S}}_{k+1}$ -$\boldsymbol{\mathcal{P}}_k$ element and $\mu = 0.1$.

k	$1/h$	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}^2_h}$	Order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$	Order	$\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0$	Order	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0$	Order	$\|p - p_h\|_0$	Order
1	4	4.6787e-01	—	1.5046e-02	—	1.2549e-01	—	4.3029e-02	—	2.4714e-02	—
	8	2.3716e-01	0.98	7.5307e-03	1.00	6.2512e-02	1.01	1.8779e-02	1.20	9.9083e-03	1.32
	16	1.1896e-01	1.00	3.7522e-03	1.01	3.0975e-02	1.01	8.9093e-03	1.08	4.5275e-03	1.13
	32	5.9528e-02	1.00	1.8747e-03	1.00	1.5423e-02	1.01	4.3814e-03	1.02	2.2002e-03	1.04
2	4	4.3240e-03	—	4.4497e-04	—	1.2330e-03	—	3.3344e-04	—	1.5870e-04	—
	8	1.2686e-03	1.77	1.1611e-04	1.94	2.2975e-04	2.42	5.8879e-05	2.50	2.6031e-05	2.61
	16	3.3791e-04	1.91	2.9412e-05	1.98	3.7033e-05	2.63	9.3965e-06	2.65	4.0887e-06	2.67
	32	8.6467e-05	1.97	7.3810e-06	1.99	6.6945e-06	2.47	1.7725e-06	2.41	8.2132e-07	2.32
3	4	9.8083e-04	—	7.5938e-05	—	1.8039e-04	—	4.7370e-05	—	2.1705e-05	—
	8	1.3462e-04	2.87	1.0074e-05	2.91	1.7161e-05	3.39	4.5104e-06	3.39	2.0692e-06	3.39
	16	1.7187e-05	2.97	1.2816e-06	2.97	1.3491e-06	3.67	3.5400e-07	3.67	1.6204e-07	3.67
	32	2.1519e-06	3.00	1.6095e-07	2.99	9.2598e-08	3.86	2.4259e-08	3.87	1.1080e-08	3.87

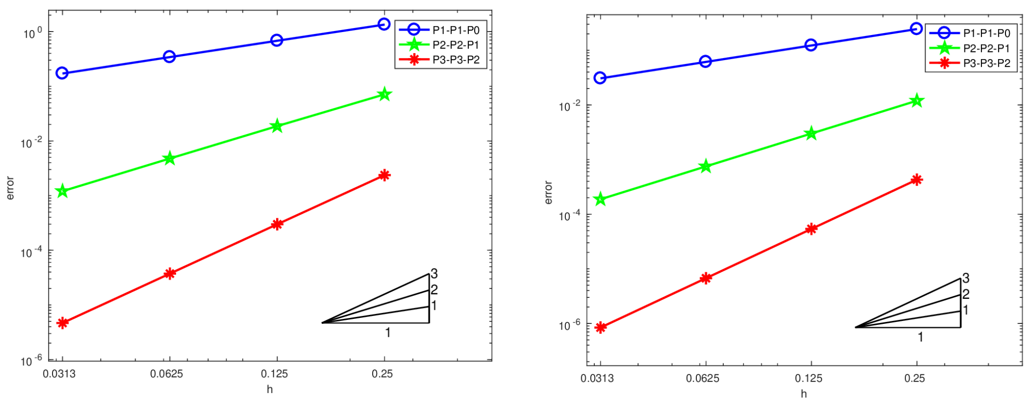
Table 5.4: Numerical errors and orders in Example 5.2 for $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$ element and $\mu = 0.01$.

k	$1/h$	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2}$	Order	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$	Order	$\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0$	Order	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0$	Order	$\|p - p_h\|_0$	Order
1	4	4.7414e-01	—	3.5176e-02	—	4.9650e-01	—	2.3358e-02	—	1.4950e-02	—
	8	2.4016e-01	0.98	1.4729e-02	1.26	1.9032e-01	1.38	6.6562e-03	1.81	3.8611e-03	1.95
	16	1.2045e-01	1.00	6.8448e-03	1.11	7.9058e-02	1.27	2.4286e-03	1.45	1.3034e-03	1.57
	32	6.0268e-02	1.00	3.3502e-03	1.03	3.6661e-02	1.11	1.0639e-03	1.19	5.4506e-04	1.26
2	4	1.2530e-03	—	5.7870e-04	—	4.0476e-03	—	9.9703e-05	—	4.1155e-05	—
	8	5.3056e-04	1.24	1.2757e-04	2.18	9.9243e-04	2.03	2.4796e-05	2.01	1.0509e-05	1.97
	16	1.6038e-04	1.73	2.9784e-05	2.10	1.6380e-04	2.60	4.0836e-06	2.60	1.7238e-06	2.61
	32	4.2734e-05	1.91	7.3942e-06	2.01	2.3479e-05	2.80	5.8877e-07	2.79	2.5114e-07	2.78
3	4	3.4223e-04	—	8.2725e-05	—	7.4805e-04	—	2.1048e-05	—	1.0468e-05	—
	8	5.1001e-05	2.75	1.0185e-05	3.02	7.9871e-05	3.23	2.0696e-06	3.35	9.3045e-07	3.49
	16	7.0224e-06	2.86	1.2840e-06	2.99	8.0826e-06	3.30	2.0662e-07	3.32	9.0994e-08	3.35
	32	9.4177e-07	2.90	1.6098e-07	3.00	6.4842e-07	3.64	1.6679e-08	3.63	7.4170e-09	3.62

Example 5.3. Consider the quasi-Newtonian Stokes flows (1.1) in a square domain $(0,1) \times (0,1)$. The right-hand side function $\boldsymbol{f}$, the exact strain rate function $\underline{\boldsymbol{t}}$, the exact stress function $\underline{\boldsymbol{\sigma}}$ and boundary condition $\boldsymbol{g}$ follow from the exact solution

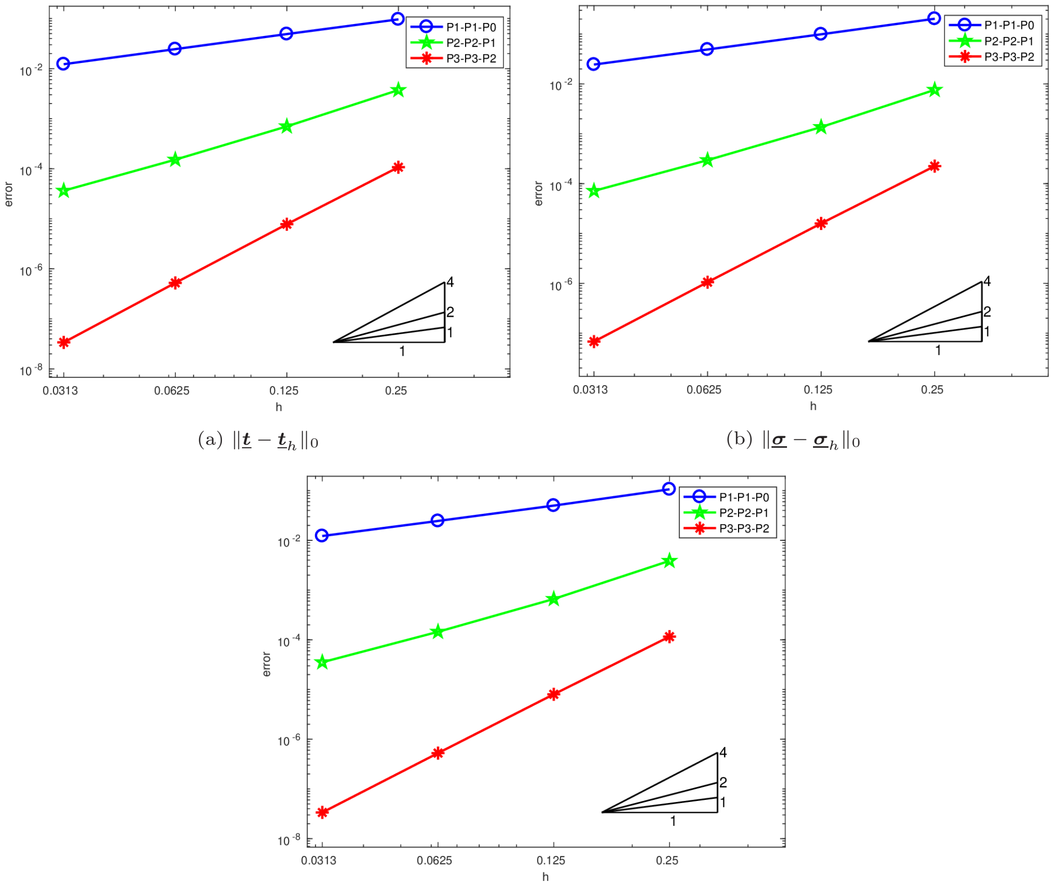
$$\begin{cases} u_1(x,y) = -\big(\sin(y) + y\cos(y)\big)e^x, \\ u_2(x,y) = y\sin(y)e^x, \\ p(x,y) = 2\sin(y)e^x - 2(e-1)\big(1-\cos(1)\big). \end{cases}$$

In this example, we also consider the Carreau's law by taking $\mu_0 = 0.5, \mu_1 = 0.5$ and $\beta = 1.5$, respectively. Similarly, we set the parameter $\kappa = 0.1$ and $h = 2^{-2}, 2^{-3}, 2^{-4}, 2^{-5}$. We compute the numerical solutions for finite element triplets $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$- $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$, $k = 0, 1, 2$ on the uniform triangulation mesh. With respected to meshsize h, the optimal convergence orders $\mathcal{O}(h^{k+1})$ of the stress in broken $\underline{\boldsymbol{H}}(\mathbf{div})$-norm and velocity in $\boldsymbol{L}^2$-norm are observed for the finite element



(a) $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\underline{\boldsymbol{\Sigma}}_h^2}$ (b) $\|\mathbf{u} - \mathbf{u}_h\|_0$

Fig. 5.1. Numerical convergence orders $\mathcal{O}(h^{k+1})$ for Example 5.3 with the finite element triplets $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$.



(a) $\|\underline{\boldsymbol{t}} - \underline{\boldsymbol{t}}_h\|_0$

(b) $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_0$

(c) $\|p - p_h\|_0$

Fig. 5.2. Numerical convergence orders $\mathcal{O}(h^{k+1})$ or $\mathcal{O}(h^{k+2})$ for Example 5.3 with the finite element triplets $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$.

triplets $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$, $k = 0, 1, 2$, see Fig. 5.1. While, Fig. 5.2 shows that the strain rate and the stress in $\underline{\boldsymbol{L}}^2$-norm, and the pressure in L^2-norm reach the optimal convergence order $\mathcal{O}(h^{k+2})$ when $k = 2$, and convergence order $\mathcal{O}(h^{k+1})$ when $k = 0, 1$, which coincides with the results in Theorems 4.2-4.3, i.e., the error estimates are sharp.

6. Summary

In this paper, the mixed discontinuous Galerkin method with symmetric strain rate and stress element triplet $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$- $\underline{\boldsymbol{\mathcal{P}}}_{k+1}^{\mathbb{S}}$-$\boldsymbol{\mathcal{P}}_k$ is constructed and studied for solving a class of quasi-Newtonian Stokes flows. The well-posedness of the MDG scheme and a priori error analysis are obtained. For any $k \geq 0$, the optimal convergence order $\mathcal{O}(h^{k+1})$ for the stress in broken $\underline{\boldsymbol{H}}(\mathbf{div})$ norm and velocity in $\boldsymbol{L}^2$ norm are proved. Furthermore, under certain assumptions, we prove the optimal $\underline{\boldsymbol{L}}^2$ error estimates of order $\mathcal{O}(h^{k+2})$ for the strain rate and the stress, and the optimal L^2 error estimate of order $\mathcal{O}(h^{k+2})$ for the pressure. Numerical examples verify the performance of the MDG scheme and confirm the theoretical results. Especially, we provide numerical evidence showing that the orders of convergence are sharp.

Acknowledgments. The work of F. Wang was partially supported by the National Natural Science Foundation of China (Grant No. 12171383) and the work of W. Yan was partially supported by the National Natural Science Foundation of China (Grant No. 11971377).

References

[1] D.N. Arnold, G. Awanou, and R. Winther, Finite elements for symmetric tensors in three dimensions, *Math. Comp.*, **77**:263 (2008), 1229–1251.

[2] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM J. Numer. Anal.*, **39**:5 (2002), 1749–1779.

[3] D.N. Arnold, J. Douglas, Jr., and C.P. Gupta, A family of higher order mixed finite element methods for plane elasticity, *Numer. Math.*, **45**:1 (1984), 1–22.

[4] W. Bao and J.W. Barrett, A priori and a posteriori error bounds for a nonconforming linear finite element approximation of a non-Newtonian flow, *RAIRO Modél. Math. Anal. Numér.*, **32**:7 (1998), 843–858.

[5] J. Baranger, K. Najib, and D. Sandri, Numerical analysis of a three-fields model for a quasi-Newtonian flow, *Comput. Methods Appl. Mech. Engrg.*, **109**:3-4 (1993), 281–292.

[6] J.W. Barrett and W.B. Liu, Finite element error analysis of a quasi-Newtonian flow obeying the Carreau or power law, *Numer. Math.*, **64**:4 (1993), 433–453.

[7] M.A. Behr, L.P. Franca, and T.E. Tezduyar, Stabilized finite element methods for the velocity-pressure-stress formulation of incompressible flows, *Comput. Methods Appl. Mech. Engrg.*, **104**:1 (1993), 31–48.

[8] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, in: *Springer Series in Computational Mathematics*, Vol. 44, Springer, 2013.

[9] F. Brezzi, J. Douglas, Jr., R. Durán, and M. Fortin, Mixed finite elements for second order elliptic problems in three variables, *Numer. Math.*, **51**:2 (1987), 237–250.

[10] F. Brezzi, J. Douglas, Jr., and L.D. Marini, Two families of mixed finite elements for second order elliptic problems, *Numer. Math.*, **47**:2 (1985), 217–235.

[11] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, in: *Springer Series in Computational Mathematics*, Vol. 15, Springer-Verlag, 1991.

[12] F. Brezzi, G. Manzini, L.D. Marini, P. Pietra, and A. Russo, Discontinuous Galerkin approximations for elliptic problems, *Numer. Methods Partial Differential Equations*, **16**:4 (2000), 365–378.

[13] R. Bustinza and G.N. Gaticam, A mixed local discontinuous Galerkin method for a class of nonlinear problems in fluid mechanics, *J. Comput. Phys.*, **207**:2 (2005), 427–456.

[14] E. Cáceres, G.N. Gatica, and F.A. Sequeira, A mixed virtual element method for quasi-Newtonian Stokes flows, *SIAM J. Numer. Anal.*, **56**:1 (2018), 317–343.

[15] Z.Q. Cai, C. Tong, P.S. Vassilevski, and C.B. Wang, Mixed finite element methods for incompressible flow: Stationary Stokes equations, *Numer. Methods Partial Differential Equations*, **26**:4 (2010), 957–978.

[16] S. Congreve, P. Houston, E. Süli, and T.P. Wihler, Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems II: Strongly monotone quasi-Newtonian flows, *IMA J. Numer. Anal.*, **33**:4 (2013), 1386–1415.

[17] Q. Du and M.D. Gunzburger, Finite-element approximations of a Ladyzhenskaya model for stationary incompressible viscous flow, *SIAM J. Numer. Anal.*, **27**:1 (1990), 1–19.

[18] V.J. Ervin, J.S. Howell, and I. Stanculescu, A dual-mixed approximation method for a three-field model of a nonlinear generalized Stokes problem, *Comput. Methods Appl. Mech. Engrg.*, **197**:33-40 (2008), 2886–2900.

[19] M. Farhloul and M. Fortin, A new mixed finite element for the Stokes and elasticity problems, *SIAM J. Numer. Anal.*, **30**:4 (1993), 971–990.

[20] G.N. Gatica, L.F. Gatica, and F.A. Sequeira, Analysis of an augmented pseudostress-based mixed formulation for a nonlinear Brinkman model of porous media flow, *Comput. Methods Appl. Mech. Engrg.*, **289** (2015), 104–130.

[21] G.N. Gatica, M. González, and S. Meddahi, A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. Part I: A priori error analysis, *Comput. Methods Appl. Mech. Engrg.*, **193**:9-11 (2004), 881–892.

[22] G.N. Gatica, M. González, and S. Meddahi, A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. Part II: A posteriori error analysis, *Comput. Methods Appl. Mech. Engrg.*, **193**:9-11 (2004), 893–911.

[23] G.N. Gatica, A. Márquez, and M.A. Sánchez, Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equations, *Comput. Methods Appl. Mech. Engrg.*, **199**:17-20 (2010), 1064–1079.

[24] G.N. Gatica, A. Márquez, and M.A. Sánchez, A priori and a posteriori error analyses of a velocity-pseudostress formulation for a class of quasi-Newtonian Stokes flows, *Comput. Methods Appl. Mech. Engrg.*, **200**:17-20 (2011), 1619–1636.

[25] G.N. Gatica, M. Munar, and F.A. Sequeira, A mixed virtual element method for a nonlinear Brinkman model of porous media flow, *Calcolo*, **55** (2018), 21.

[26] G.N. Gatica and F.A. Sequeira, Analysis of an augmented HDG method for a class of quasi-Newtonian Stokes flows, *J. Sci. Comput.*, **65**:3 (2015), 1270–1308.

[27] G.N. Gatica and F.A. Sequeira, A priori and a posteriori error analyses of an augmented HDG method for a class of quasi-Newtonian Stokes flows, *J. Sci. Comput.*, **69**:3 (2016), 1192–1250.

[28] S.H. Gong, S.N. Wu, and J.C. Xu, New hybridized mixed methods for linear elasticity and optimal multilevel solvers, *Numer. Math.*, **141**:2 (2019), 569–604.

[29] J. Gopalakrishnan and J. Guzman, Symmetric nonconforming mixed finite elements for linear elasticity, *SIAM J. Numer. Anal.*, **49**:4 (2011), 1504–1520.

[30] C. Han, *Multiphase Flow in Polymer Processing*, Academic Press, 1981.

[31] J.S. Howell, Dual-mixed finite element approximation of Stokes and nonlinear Stokes problems using trace-free velocity gradients, *J. Comput. Appl. Math.*, **231**:2 (2009), 780–792.

[32] J. Hu, Finite element approximations of symmetric tensors on simplicial grids in $\mathbb{R}^n$: The higher order case, *J. Comput. Math.*, **33**:3 (2015), 283–296.

[33] J. Hu and S.Y. Zhang, A family of conforming mixed finite elements for linear elasticity on triangular grids, *arXiv:1406.7457*, 2014.

[34] J. Hu and S.Y. Zhang, A family of symmetric mixed finite elements for linear elasticity on tetrahedral grids, *Sci. China Math.*, **58**:2 (2015), 297–307.

[35] O.A. Ladyzhenskaja, *New Equations for the Description of the Viscous Incompressible Fluids and Solvability in the Large of the Boundary Value Problems for Them*, in: *Boundary Value Problems of Mathematical Physics*, Vol. V, American Mathematical Society, 1970.

[36] H.P. Langtangen and A. Logg, *Solving PDEs in Python*, in: *Simula Springer Briefs on Computing*, Vol. 3, Springer, 2016.

[37] A.F.D. Loula and J.N.C. Guerreiro, Finite element analysis of nonlinear creeping flows, *Comput. Methods Appl. Mech. Engrg.*, **79**:1 (1990), 87–109.

[38] Y.X. Qian, S.N. Wu, and F. Wang, A mixed discontinuous Galerkin method with symmetric stress for Brinkman problem based on the velocity-pseudostress formulation, *Comput. Methods Appl. Mech. Engrg.*, **368** (2020), 113177.

[39] D. Sandri, Sur l'approximation numérique des écoulements quasi-newtoniens dont la viscosité suit la loi puissance ou la loi de Carreau, *RAIRO Modél. Math. Anal. Numér.*, **27**:2 (1993), 131–155.

[40] B. Scheurer, Existence et approximation de points selles pour certains problèmes non linéaires, *RAIRO Anal. Numér.*, **11**:4 (1997), 369–400.

[41] F. Wang, S.N. Wu, and J.C. Xu, A mixed discontinuous Galerkin method for linear elasticity with strongly imposed symmetry, *J. Sci. Comput.*, **83** (2020), 2.

[42] S.N. Wu, S.H. Gong, and J.C. Xu, Interior penalty mixed finite element methods of any order in any dimension for linear elasticity with strongly symmetric stress tensor, *Math. Models Methods Appl. Sci.*, **27**:14 (2017), 2711–2743.

[43] X.B. Zheng, G. Chen, and X.P. Xie, A divergence-free weak Galerkin method for quasi-Newtonian Stokes flows, *Sci. China Math.*, **60**:8 (2017), 1515–1528.