Penalty finite element approximations of the stationary power-law Stokes problem

Lew Lefton [*] and Dongming Wei [†]

Abstract — Finite element approximations of the stationary power-law Stokes problem using penalty formulation are considered. A priori error estimates under appropriate smoothness assumptions on the solutions are established without assuming a discrete version of the BB condition. Numerical solutions are presented by implementing a nonlinear conjugate gradient method.

Keywords: Power-law flows, penalty method, stationary Stokes problem, non-Newtonian flows, finite element method, convergence and error estimates, BB condition

1. INTRODUCTION

Consider the following stationary power-law Stokes problem:

$$
\begin{aligned}
-k\nabla \cdot (|\mathbf{D}(\mathbf{u})|^{r-2}\mathbf{D}(\mathbf{u})) + \nabla p &= \mathbf{f} && \text{in } \Omega, \\
\nabla \cdot \mathbf{u} &= 0 && \text{in } \Omega, \\
\mathbf{u} &= \mathbf{0} && \text{on } \partial\Omega.
\end{aligned}
\tag{1.1}
$$

In the above, Ω is a bounded convex domain in $\mathbb{R}^d$, $d = 2$ or 3, $\partial\Omega$ the boundary of Ω, $\mathbf{u} = (u_1(\mathbf{x}),...,u_d(\mathbf{x}))$ the velocity of a fluid particle located at $\mathbf{x} = (x_1,...,x_d) \in \Omega$, $\nabla \cdot \mathbf{u} = \sum_{i=1}^{d} \frac{\partial u_i}{\partial x_i}$ the divergence of $\mathbf{u}$, ∇p the gradient of the fluid pressure p, $\mathbf{f} = \mathbf{f}(\mathbf{x})$ the body force, the constant k a measure of the consistency of the fluid, and $1 < r < \infty$. The rate of deformation tensor $\mathbf{D}(\mathbf{u}) \in \mathbb{R}^d \times \mathbb{R}^d$ is the symmetric gradient of $\mathbf{u}$ with components $d_{ij}(\mathbf{u}) = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$ and $|D(\mathbf{u})| = [\sum_{i,j=1}^{d} (d_{ij}(\mathbf{u}))^2]^{\frac{1}{2}}$. A second problem for stationary power-law Stokesian flows is to consider the velocity field $\mathbf{u}$ as the minimizer of an appropriate energy functional in the space of divergence free

[*]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30032, llefton@math.gatech.edu

[†]Department of Mathematics, University of New Orleans, New Orleans, Louisiana 70148, dwei@uno.edu

vector fields, that is, to solve the problem:

$$\text{Minimize } J(\mathbf{u}) = \frac{k}{r} \int_\Omega |\mathbf{D}(\mathbf{u})|^r \, d\mathbf{x} - \int_\Omega \mathbf{f} \cdot \mathbf{u} \, d\mathbf{x},$$

$$\text{subject to } \mathbf{u} \in \mathbf{X} = \left\{ \mathbf{u} \in \mathbf{W}_0^{1,r}(\Omega) : \nabla \cdot \mathbf{u} = 0 \right\}. \tag{1.2}$$

Problem (1.1) is called the mixed formulation of the stationary power-law Stokesian flow which has a weak form written down precisely in (2.6). Problem (1.2) is referred to as the *variational formulation* of the stationary power-law Stokesian flows. The Euler-Lagrange equation corresponding to this minimization problem is solved for velocity in a subspace of divergence free vector fields. It is well known that the connection between these two formulations comes from a technical inf-sup condition which is frequently called the *BB condition* named after Brezzi [19] and Babuška [7]. The *BB condition* is also called the *LBB condition* for it is related to the work of Ladyzhenskaya [42]. This condition can be stated as the following: $\exists \beta > 0$ such that

$$\inf_{q \in L_0^{r'}(\Omega)} \sup_{\mathbf{v} \in \mathbf{W}_0^{1,r}(\Omega)} \frac{\langle \nabla \cdot \mathbf{v}, q \rangle}{||q||_{0,r'} ||\mathbf{v}||_{1,r}} \geqslant \beta. \tag{1.3}$$

When (1.3) holds, the variational formulation and the mixed weak formulation are equivalent in the sense that $\mathbf{u}$ is a solution of the variational formulation if and only if $(\mathbf{u}, p)$ is a solution of the mixed weak formulation with p being solved in terms of $\mathbf{u}$ using the inverse of the divergence operator.

The *BB* condition holds for all dimensions $d > 1$ and for the full range $1 < r < \infty$ in bounded Lipschitz domains in $\mathbb{R}^d$ as stated in Amrouche and Girault [6]. Therefore, the mixed weak formulation is well-posed and it is equivalent to the variational formulation for bounded convex domains in $\mathbb{R}^d$, since such domains are Lipschitz.

Since, in the variational formulation, $\mathbf{u}$ is defined as the minimum of a convex functional on a separable Banach space, (1.2) always has a unique solution $\mathbf{u}$ for any $1 < r < \infty$, and for any bounded domain Ω in $\mathbb{R}^d$, whether or not the *BB* condition holds. However, the pressure p is not necessarily well defined, so that results involving the pressure function p typically require the additional assumption of an *BB* condition. It is not known if (1.3) holds in non-Lipschitz bounded domains.

The two formulations discussed above for power-law Stokes flows are both useful for the numerical analysis of the problem. Finite element analysis of power-law Stokes flows using the mixed weak formulation has been studied by several authors, for example, [8], [9], [10], and [11]. Baranger, et al, [8], [9], obtain error estimates which require a discrete version of the *BB* condition, namely, $\exists \beta > 0$ independent of h, such that

$$\inf_{q \in M_0^h(\Omega)} \sup_{\mathbf{v} \in \mathbf{S}_0^h(\Omega)} \frac{\langle \nabla \cdot \mathbf{v}, q \rangle}{||q||_{0,r'} ||\mathbf{v}||_{1,r}} \geqslant \beta, \tag{1.4}$$

where $M_0^h(\Omega) \times \mathbf{S}_0^h(\Omega)$ is the corresponding finite element space for $L_0^{r'}(\Omega) \times \mathbf{W}_0^{1,r}(\Omega)$. They show that for $1 < r \leqslant 2$ and $d = 2$, if the *BB* condition (1.3) holds,

then

$$||\mathbf{u} - \mathbf{u}_h||_{1,r} \leqslant \inf_{(\mathbf{v}_h,q_h) \in \mathbf{S}_0^h(\Omega) \times M_0^h(\Omega)} \left(C_1 ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{r}{2}} + C_2 ||p - q_h||_{0,r'}^{\frac{1}{2}} \right).$$

If, in addition to (1.3), the discrete version of the *BB* condition (1.4) holds, then,

$$||p - p_h||_{0,r'} \leqslant \inf_{(\mathbf{v}_h,q_h) \in \mathbf{S}_0^h(\Omega) \times M_0^h(\Omega)} \left(C_1 ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{r-1} + C_2 ||p - q_h||_{0,r'} \right),$$

where C_1 and C_2 depend on β. Here we are writing $(\mathbf{u}, p)$ for the unique solution of the mixed formulation of (1.1) as written in (2.6), and $(\mathbf{u}_h, p_h)$ for the corresponding finite element approximation to $(\mathbf{u}, p)$ using the corresponding discrete mixed formulation. Barrett and Liu obtained the following similar results:

$$||\mathbf{u} - \mathbf{u}_h||_{1,r} \leqslant \inf_{(\mathbf{v}_h,q_h) \in \mathbf{S}_0^h(\Omega) \times M_0^h(\Omega)} \left(C_1 ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{r}{2}} + C_2 ||p - q_h||_{0,r'} \right)$$

and

$$||p - p_h||_{0,r'} \leqslant \inf_{(\mathbf{v}_h,q_h) \in \mathbf{S}_0^h(\Omega) \times M_0^h(\Omega)} \left(C_1 ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{r-1} + C_2 ||p - q_h||_{0,r'} \right)$$

under the same conditions. In addition, Barrett and Liu [11] also extended the results to the case when $2 \leqslant r < \infty$.

Finite element analysis using the direct variational formulation (1.2) requires one to solve a constrained minimization problem and construct finite element spaces with divergence free or approximately divergence free interpolation functions. Thus, the variational formulation and its associated constrained minimization problem is more difficult for both analysis and practical numerical approximation. Divergence free elements have been studied extensively and it is not intended here to comment on the results related to this approach. A natural way to overcome this difficulty is to introduce a penalty functional that eliminates the constraint. Although, error analysis and computer implementation becomes relatively simple, the resultant discrete system from the penalty requires the mesh size to be proportional to the penalty parameter for convergence. Therefore, a finite element mesh has to be very fine for the incompressibility condition to be close to being satisfied. However, with the help of the techniques of reduced integration and the high speed computers, this deficiency can be degraded and penalty method for fluid flow problems is still a popular choice among scientists and engineers, see, e.g., [34], [35], [43], [44], [55], and [56] for applications of the penalty method to power-law flow problems.

The first use of the penalty function method in conjunction with the finite element method is due to Babǔska [7]. The method was quickly adopted as a standard tool for the finite element analysis of viscous, incompressible fluid flows [66]. Extensive studies of the penalty method applied to Newtonian fluid flow problems, both experimentally and mathematically, have been conducted from the late seventies to the present day. Here we cite only several important articles in this area:

[13], [14], [31], [32], [33], [39], [40], [49], [50], [51], [53], [54], and [66]. A very general mathematical analysis of the penalty method applied to nonlinear problems including a class of non-Newtonian fluid flow problems was presented by Oden [51]. His work provides some important convergence results. For a given penalty parameter $\varepsilon > 0$, the penalty formulation requires the unconstrained minimization of the nonlinear convex functional

$$J_\varepsilon(\mathbf{u}) = J(\mathbf{u}) + \frac{1}{r\varepsilon} \int_\Omega |\nabla \cdot \mathbf{u}|^r d\mathbf{x}$$

over the Sobolev space $\mathbf{W}_0^{1,r}(\Omega)$. The corresponding pressure p^ε is defined in terms of the minimizer $\mathbf{u}^\varepsilon$. It is known that the penalty approximation $\mathbf{u}^\varepsilon$ of the unconstrained minimization problem $\min\{J_\varepsilon(\mathbf{u}) : \mathbf{u} \in \mathbf{W}_0^{1,r}(\Omega)\}$ converges to the true solution $\mathbf{u}$ of $\min\{J(\mathbf{u}) : \mathbf{u} \in X\}$ as $\varepsilon \to 0$ for any $1 < r < \infty$ without assuming that the domain Ω is convex (see, e.g., [51], [65], [45]). This convergence result is only for the velocity field since the pressure may be undefined. However, because of the more general variational setting, this result establishes the validity of a penalty approximation even when the *BB* condition fails to hold. This is a convergence result, not an error estimate, but it doesn't require the *BB* condition (1.3).

When the *BB* condition holds, it has been shown in [45] that the error estimates for the velocity field are given by $\|\mathbf{u} - \mathbf{u}^\varepsilon\|_{1,r} = O(\varepsilon^{g_1(r)})$, where $g_1(r) = \frac{1}{(r-1)(3-r)}$ for $1 < r \leqslant 2$, and $g_1(r) = \frac{1}{(r-1)^2}$ for $2 \leqslant r < \infty$. Let $\varphi(z) = |z|^{r-2}z, z \in \mathbb{R}$ and let $p^\varepsilon = c - \frac{1}{\varepsilon}\varphi(\nabla \cdot \mathbf{u}^\varepsilon)$, where $c = \int_\Omega \frac{1}{\varepsilon|\Omega|}\varphi(\nabla \cdot \mathbf{u}^\varepsilon)d\mathbf{x}$. The corresponding error estimates for the pressure are given by $\|p - p^\varepsilon\|_{0,r'} = O(\varepsilon^{g_2(r)})$ where $g_2(r) = \frac{1}{(3-r)}$ for $1 < r \leqslant 2$ and $g_2(r) = \frac{1}{(r-1)^2}$ for $2 \leqslant r < \infty$. These rates of convergence reduce to known results $\|\mathbf{u} - \mathbf{u}^\varepsilon\|_{1,2} + \|p - p^\varepsilon\|_{0,2} = O(\varepsilon)$ for the Newtonian case $r = 2$ as discussed in [13], [14], [37], [49], [53], and [59]. The above error estimates are generalized for the corresponding power-law Navier-Stokes problem in the case $r \neq 2$, see [62] for details.

The penalty finite element numerical scheme for approximating (1.2) is to fix a small positive ε, and solve for the minimizer $\mathbf{u}_h^\varepsilon$ of the problem $\min\{J_\varepsilon(\mathbf{u}) : \mathbf{u} \in \mathbf{S}_0^h(\Omega)\}$. The main result of this work is the following:

Suppose that Ω is a bounded convex domain in $\mathbb{R}^d$ and $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$. For $0 < \varepsilon \leqslant 1$, let $\mathbf{u}$ be the unique minimizer of the functional J over $\mathbf{W}_0^{1,r}(\Omega)$, and let $\mathbf{u}_h^\varepsilon$ be the unique minimizer of $J_\varepsilon(\mathbf{u})$ over $\mathbf{S}_0^h(\Omega)$. Then, there exists a positive constant C, which depends only on Ω, $\mathbf{f}$, and r, such that $\forall \mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$

$$C\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(3-r)}}\|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{(3-r)}} + \|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{2}} + \varepsilon^{\frac{1}{2r}},$$

for $1 < r \leqslant 2$, and

$$C\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(r-1)}}\|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{(r-1)}} + \|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{r}} + \varepsilon^{\frac{1}{r^2}},$$

for $2 \leqslant r < \infty$. Furthermore, let $\mathbf{V} = \{\mathbf{v} \in C_0^\infty(\Omega) : \nabla \cdot \mathbf{v} = 0\}$, and suppose $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ satisfies $\forall \varphi \in \mathbf{V}, \langle \mathbf{f}, \varphi \rangle = 0$. Then $\forall \mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$

$$C\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,r} \leqslant (\frac{1}{\varepsilon})^{\frac{1}{2(3-r)}} \|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{(3-r)}} + \|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{2}} + \varepsilon^{\frac{1}{2(r-1)}}$$

for $1 < r \leqslant 2$, and

$$C\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,r} \leqslant (\frac{1}{\varepsilon})^{\frac{1}{r(r-1)^2}} \|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{r-1}} + \|\mathbf{u} - \mathbf{v}_h\|_{1,r}^{\frac{1}{r}} + \varepsilon^{\frac{1}{r(r-1)}}$$

for $2 \leqslant r < \infty$.

The above results are important, in particular, if $r = 2$, $1 < \alpha < \infty$, $\ell \geqslant 2$ is an integer, Ω is a bounded and open set of $\mathbb{R}^d$ of class C^ℓ, and $\mathbf{f} \in \mathbf{W}^{\ell-2,\alpha}(\Omega)$, then there exists a positive constant C, independent of ε, such that $\|\mathbf{u}\|_{\ell,\alpha} \leqslant C$, and

$$\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,2} \leqslant C\left(\varepsilon^{-\frac{1}{2}} h^{\ell-1+d(\frac{1}{2}-\frac{1}{\alpha})} + h^{\frac{1}{2}[\ell-1+d(\frac{1}{2}-\frac{1}{\alpha})]} + \varepsilon^{\frac{1}{2}}\right),$$

with C independent of ε. If $\ell = 2$ and we choose $\varepsilon = h$, we then have the optimal error estimate in velocity for linear elements:

$$\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,2} \leqslant C(h^{\frac{1}{2}} + h^{\frac{1}{2}(1+d(\frac{1}{2}-\frac{1}{\alpha}))}).$$

This reduces to the same rate of $O(h^{\frac{1}{2}})$ found in the paper of Reddy [53] for $\alpha = 2$ under similar conditions. It is shown by Falk and King [33] that for this "simple" penalty method, if $r = \alpha = 2$, $\varepsilon = h$, and Ω is an open bounded set of class C^ℓ, then

$$\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,2} = O(h^{\frac{2}{3}(\ell-1)}),$$

with appropriate choice of the finite element space $\mathbf{S}_0^h(\Omega)$. They also show a higher order rate of accuracy than that given above can be obtained by using extrapolation techniques (see [33]). A better rate of $O(h)$ is shown by Falk [31] by using a modified penalty term. Oden, Kikuchi and Song [49] also showed that

$$\|\mathbf{u} - \mathbf{u}^h\|_{1,2} + \|p - p^h\|_{0,2} \leqslant C(\|\mathbf{u}\|_{2,2} + \|p\|_{1,2})h$$

where $(\mathbf{u}^h, p^h)$ is obtained by using a mixed finite element approximation as the limit, as $\varepsilon \to 0$, with a certain Gaussian integration rule for some domains. For $r \neq 2$, an uniform bound on $\|\mathbf{u}\|_{\ell,\alpha}$ for $\ell \geqslant 2$ is not available due to lack of regularity results for the degenerate nonlinear elliptic system (1.1). Regularity of solutions to nonlinear elliptic systems is an active area of research. Some related results found in [21], [25], and [60] indicate that the best possible bound for $\mathbf{u}$ is in $C^{1,\mu}(\bar{\Omega})$ with $0 < \mu < 1$. If this is the case, then $\|\mathbf{u}\|_{1,\infty} \leqslant C$, where C is independent of ε and we have

$$\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,r} = O(h^{\frac{d}{r}-\frac{1}{2}}).$$

The results of this work are useful, because although the continuous *BB* condition (1.3) must hold, to estimate $\|\mathbf{u} - \mathbf{u}^\varepsilon\|_{1,r}$ and $\|p - p^\varepsilon\|_{0,r'}$, the corresponding discrete *BB* condition (1.4) is not required to obtained convergence and error estimates of $\|\mathbf{u} - \mathbf{u}_h^\varepsilon\|_{1,r}$ and $\|p - p_h^\varepsilon\|_{0,r'}$ and therefore simple finite elements can be used. This is an important feature of the unconstrained penalty minimization formulation which makes it convenient for error analysis and numerical implementation. In contrast, the mixed weak formulation requires the solution of a system of nonlinear equations and a discrete *BB* condition.

This work provides a mathematical finite element analysis of the penalty method applied to the power-law Stokes problem in the variational formulation which is a generalization of the analysis available in the literature for Newtonian Stokes flows. Penalty method and reduced integration for the analysis of Newtonian Stokes flows has been studied extensively by Oden, et al. See [49] and [50] for further references. Numerical experiments have been performed on power-law flow problems using the penalty method in the engineering literature (see, e.g., [43], [44], [53], [55], and [66]). Since pressure p must then be calculated from the computed velocity field $\mathbf{u}$, the accuracy of the pressure is lower than that of the velocity as shown in our error estimates. Note that the penalty term $\frac{1}{2\varepsilon} \int_\Omega |\nabla \cdot \mathbf{u}|^2 d\mathbf{x}$ was used in Reddy and Padhye [55] instead of $\frac{1}{r\varepsilon} \int_\Omega |\nabla \cdot \mathbf{u}|^r d\mathbf{x}$ and reasonable numerical results were obtained. The accuracy and convergence of the penalty solution is sensitive to the values of the penalty parameter ε as well as the finite element mesh size h and the numerical integration of the penalty term.

In Section 2, notations and previous results are reviewed including error estimates in terms of the penalty parameter ε. In Section 3, fully discrete error estimates in terms of the penalty parameter ε and the finite element mesh size h are established. In Section 4, a nonlinear conjugate gradient method is described and implemented to obtained numerical solutions for a cavity problem by using the penalty formulation. The numerical results support the theoretical analysis.

2. NOTATIONS AND PRELIMINARIES

The following are standard notations. Let $L^r(\Omega)$ for $1 < r < \infty$ be the space of real scalar functions defined on Ω whose r^{th} power is absolutely integrable with respect to Lebesgue measure $d\mathbf{x} = dx_1 \ldots dx_d$. This is a Banach space with the norm $\|u\|_{0,r} = (\int_\Omega |u(\mathbf{x})|^r d\mathbf{x})^{1/r}$. The Sobolev space $W^{k,r}(\Omega)$ is the space of functions in $L^r(\Omega)$ with distributional derivatives up to order k also in $L^r(\Omega)$. The norm for this space is $\|u\|_{k,r} = (\int_\Omega \sum_{|j| \leqslant k} |D^j u(\mathbf{x})|^r d\mathbf{x})^{1/r}$, where we use the standard multi-index notation. That is, for $j = (j_1 \ldots j_d) \in \mathbb{N}^d$, where $\mathbb{N}$ is the set of natural numbers, define $|j| = \sum_{i=1}^{d} j_i$ and write the partial derivative $D^j u(\mathbf{x}) = \frac{\partial^{|j|} u}{\partial^{j_1} x_1 \ldots \partial^{j_d} x_d}$. The closure of $C_0^\infty(\Omega)$ in $W^{k,r}(\Omega)$ is denoted by $W_0^{k,r}(\Omega)$. For systems of equations, we need the product spaces defined by $\mathbf{L}^r(\Omega) = [L^r(\Omega)]^d$, $\mathbf{W}^{k,r}(\Omega) = [W^{k,r}(\Omega)]^d$, and $\mathbf{W}_0^{k,r}(\Omega) = [W_0^{k,r}(\Omega)]^d$. The norm for $\mathbf{v} = (v_1, \ldots, v_d) \in \mathbf{L}^r(\Omega)$ is $\|\mathbf{v}\|_{0,r} = (\int_\Omega \sum_{i=1}^{d} |v_i|^r d\mathbf{x})^{1/r}$.

For $\mathbf{v} \in \mathbf{W}^{k,r}(\Omega)$ we have norm $\|\mathbf{v}\|_{k,r} = (\int_\Omega \sum_{i=1}^d \sum_{|j| \leqslant k} |D^j v_i|^r d\mathbf{x})^{1/r}$. It is well known, [1], that the seminorm $|\mathbf{v}|_{k,r} = (\int_\Omega (\sum_{i=1}^d \sum_{|j|=k} |D^j v_i|^r d\mathbf{x})^{1/r}$ is equivalent to $\|\mathbf{v}\|_{k,r}$ for $\mathbf{v} \in \mathbf{W}_0^{k,r}(\Omega)$. In addition, by Korn's inequality [1], the norm $\|\mathbf{D}(\mathbf{u})\|_{0,r} = (\int_\Omega \sum_{i,j=1}^d |D_{ij}(\mathbf{u})|^r d\mathbf{x})^{1/r}$ is equivalent to $\|\mathbf{u}\|_{1,r}$ in $\mathbf{W}_0^{1,r}(\Omega)$. For $1 < r < \infty$ let r' satisfy $\frac{1}{r} + \frac{1}{r'} = 1$, which is equivalent to $r' = \frac{r}{r-1}$. Let $\|\cdot\|_{-1,r'}$ denote the norm on $\mathbf{W}^{-1,r'}(\Omega)$, which is the dual space of $\mathbf{W}_0^{1,r}(\Omega)$. Let $\varphi(\mathbf{x}) = |\mathbf{x}|^{r-2}\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^d$. Note that $\varphi^{-1}(\mathbf{x}) = |\mathbf{x}|^{r'-2}\mathbf{x}$. Finally, let $L_0^r(\Omega) = \{q \in L^r(\Omega) : \int_\Omega q \, d\mathbf{x} = 0\}$.

The following inequalities hold for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$; the constant $C > 0$ is independent of $\mathbf{x}$ and $\mathbf{y}$.

$$\begin{aligned}
&|\mathbf{x} - \mathbf{y}|^2 \leqslant C(\varphi(\mathbf{x}) - \varphi(\mathbf{y})) \cdot (\mathbf{x} - \mathbf{y})(|\mathbf{x}| + |\mathbf{y}|)^{2-r}, \\
&|\varphi(\mathbf{x}) - \varphi(\mathbf{y})| \leqslant C|\mathbf{x} - \mathbf{y}|^{r-1}, \text{ for } 1 < r < 2;
\end{aligned} \tag{2.1}$$

$$\begin{aligned}
&|\mathbf{x} - \mathbf{y}|^r \leqslant C\,(\varphi(\mathbf{x}) - \varphi(\mathbf{y})) \cdot (\mathbf{x} - \mathbf{y}), \\
&|\varphi(\mathbf{x}) - \varphi(\mathbf{y})| \leqslant C|\mathbf{x} - \mathbf{y}|(|\mathbf{x}| + |\mathbf{y}|)^{r-2}, \text{ for } 2 \leqslant r < \infty.
\end{aligned} \tag{2.2}$$

They were proved for the case $d = 2$ by Glowinski and Marroco [38] and were generalized by Barrett and Liu [11], see also Chow [22]. A simple proof for general d is shown in the book by DiBenedetto [25].

The Euler-Lagrange equation associated with problem (1.2) is

$$\langle A(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{f}, \mathbf{v} \rangle \qquad \forall \, \mathbf{v} \in \mathbf{X}, \tag{2.3}$$

where $A : \mathbf{W}_0^{1,r}(\Omega) \to \mathbf{W}^{-1,r'}(\Omega)$ is defined by

$$\langle A(\mathbf{u}), \mathbf{v} \rangle = k \int_\Omega |\mathbf{D}(\mathbf{u})|^{r-2}\mathbf{D}(\mathbf{u}) : \mathbf{D}(\mathbf{v}) dx \qquad \forall \mathbf{v} \in \mathbf{W}_0^{1,r}(\Omega),$$

and $\mathbf{X} = \left\{ \mathbf{u} \in \mathbf{W}_0^{1,r}(\Omega) : \nabla \cdot \mathbf{u} = 0 \right\}$.

Let us use $\langle \cdot, \cdot \rangle$ to denote the duality pairing between $\mathbf{W}_0^{1,r}(\Omega)$ and $\mathbf{W}^{-1,r'}(\Omega)$ as well as between $\mathbf{L}^r(\Omega)$ and $\mathbf{L}^{r'}(\Omega)$. In particular, we only need to assume $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ for this formulation.

It is well-know that due to (2.1) and (2.2), the following holds

$$\begin{aligned}
&\|\mathbf{u} - \mathbf{v}\|_{1,r}^2 \leqslant C\langle A(\mathbf{u}) - A(\mathbf{v}), \mathbf{u} - \mathbf{v} \rangle(\|\mathbf{u}\|_{1,r} + \|\mathbf{v}\|_{1,r})^{2-r}, \\
&\|A(\mathbf{u}) - A(\mathbf{v})\|_{-1,r'} \leqslant C\|\mathbf{u} - \mathbf{v}\|_{1,r}^{r-1}, \text{ for } 1 < r \leqslant 2;
\end{aligned} \tag{2.4}$$

$$\begin{aligned}
&\|\mathbf{u} - \mathbf{v}\|_{1,r}^r \leqslant C\langle A(\mathbf{u}) - A(\mathbf{v}), \mathbf{u} - \mathbf{v} \rangle, \\
&\|A(\mathbf{u}) - A(\mathbf{v})\|_{-1,r'} \leqslant C\|\mathbf{u} - \mathbf{v}\|_{1,r}(\|\mathbf{u}\|_{1,r} + \|\mathbf{v}\|_{1,r})^{r-2}, \text{ for } 2 \leqslant r < \infty,
\end{aligned} \tag{2.5}$$

where $C > 0$ denotes a generic constant independent of $\mathbf{u}$ and $\mathbf{v}$. From (2.4) and (2.5), A can be shown to be a bounded, monotone, coercive, potential operator on

$\mathbf{X} = \{\mathbf{u} \in \mathbf{W}_0^{1,r}(\Omega) : \nabla \cdot \mathbf{u} = 0\}$. It follows that $J : \mathbf{X} \to \mathbb{R}$ is a continuous, strictly convex functional on $\mathbf{X}$ and

$$\lim_{||\mathbf{u}||_{1,r} \to \infty} J(\mathbf{u}) = \infty.$$

It is standard to show that (2.3) and (1.2) are equivalent. Note that in (2.3), unlike (1.1), pressure p does not appear.

Let $b(p, \mathbf{v}) = \int_\Omega p \nabla \cdot \mathbf{v} \, d\mathbf{x}$ be the bilinear form defined on $L_0^{r'}(\Omega) \times \mathbf{W}_0^{1,r}(\Omega)$ and let $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ then the mixed weak formulation of (1.1) is the problem of finding $(\mathbf{u}, p) \in \mathbf{W}_0^{1,r}(\Omega) \times L_0^{r'}(\Omega)$ such that

$$
\begin{aligned}
\langle A(\mathbf{u}), \mathbf{v} \rangle - b(p, \mathbf{v}) = \langle \mathbf{f}, \mathbf{v} \rangle \qquad & \forall \, \mathbf{v} \in \mathbf{W}_0^{1,r}(\Omega), \\
b(q, \mathbf{u}) = 0 \qquad & \forall \, q \in L_0^{r'}(\Omega).
\end{aligned}
\tag{2.6}
$$

The existence and uniqueness of solutions of (2.3) and (2.6) was studied by Baranger and Najib [8] and Barrett and Liu [11].

Using the theory of Baranger and Najib [8] and the *BB* condition on bounded convex domains in $\mathbb{R}^d$ stated by Amrouche and Girault [6], the following Theorem holds as a simple consequence

Theorem 2.1. *Let $1 < r < \infty$ and suppose that $\Omega \subset \mathbb{R}^d, d > 1$ is a bounded convex domain. Then* (2.6) *has a unique solution if and only if* (2.3) *has a solution and the divergence operator $\nabla \cdot$ is surjective.*

The following Lemma is due to (2.4), (2.5) and (1.3) as shown in Lefton and Wei [45].

Lemma 2.1. *Let Ω be a bounded convex domain in $\mathbb{R}^d, d > 1$ and let $(\mathbf{u}, p)$ be the solution of* (2.6). *Then $||\mathbf{u}||_{1,r} \leqslant C$ and $||p||_{0,r'} \leqslant C$, where $C > 0$ depends only on r, Ω and $\mathbf{f}$.*

3. THE PENALTY FORMULATION FOR THE STOKES PROBLEM

Let ε be a positive number and consider the following functional

$$J_\varepsilon(\mathbf{u}) = J(\mathbf{u}) + \frac{1}{r\varepsilon} \int_\Omega |\nabla \cdot \mathbf{u}|^r \, d\mathbf{x},$$

where $J(\mathbf{u})$ is defined in (1.2). The minimizer $\mathbf{u}$ of $J_\varepsilon(\mathbf{u})$ over $\mathbf{W}_0^{1,r}(\Omega)$ satisfies the corresponding Euler-Lagrange equation

$$\langle A(\mathbf{u}), \mathbf{v} \rangle + \frac{1}{\varepsilon} \langle \varphi(\nabla \cdot \mathbf{u}), \nabla \cdot \mathbf{v} \rangle = \langle \mathbf{f}, \mathbf{v} \rangle \qquad \forall \, \mathbf{v} \in \mathbf{W}_0^{1,r}(\Omega). \tag{3.1}$$

Note that $J_\varepsilon(\mathbf{u})$ is strictly convex and the operator $T_\varepsilon : \mathbf{W}_0^{1,r}(\Omega) \to \mathbf{W}^{-1,r'}(\Omega)$ defined by $T_\varepsilon(\mathbf{u}) = J'_\varepsilon(\mathbf{u})$ is a bounded, coercive, and strictly monotone operator on $\mathbf{W}_0^{1,r}(\Omega)$, since φ satisfies (2.1) and (2.2). Therefore, by the standard theory of monotone operators, (3.1) has a solution $\mathbf{u}^\varepsilon$ which is the unique solution of

$$\min_{\mathbf{u}\in\mathbf{W}_0^{1,r}(\Omega)} J_\varepsilon(\mathbf{u}). \qquad (3.2)$$

The following two Theorems are obtained in [45]:

Theorem 3.1. *Suppose that Ω is a bounded domain in $\mathbb{R}^d, d > 1$. Let $\mathbf{u}$ be the solution of (1.2) and, for a fixed $\varepsilon > 0$, let $\mathbf{u}^\varepsilon$ be the solution of (3.2). Then $\mathbf{u}^\varepsilon$ converges strongly to $\mathbf{u}$ in $\mathbf{W}_0^{1,r}(\Omega)$ as $\varepsilon \to 0$, and $\exists\, C$ independent of ε such that $||\mathbf{u}^\varepsilon||_{1,r} \leqslant C$ and $||\nabla \cdot \mathbf{u}^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r}}$. Furthermore, if (1.3) holds, then $||\nabla \cdot \mathbf{u}^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r-1}}$.*

Theorem 3.2. *Suppose that Ω is a bounded convex domain in $\mathbb{R}^d, d > 1$. Let $(\mathbf{u}, p)$ be the unique solution of (2.6) and $\mathbf{u}^\varepsilon$ be a solution of (3.2). Let $p^\varepsilon = c - \frac{1}{\varepsilon}\varphi(\nabla \cdot \mathbf{u}^\varepsilon)$, where $c = \frac{1}{\varepsilon|\Omega|}\int_\Omega \varphi(\nabla \cdot \mathbf{u}^\varepsilon)d\mathbf{x}$. Then there exists $C > 0$ which depends only on r, Ω, and $\mathbf{f}$ such that*

$$||\mathbf{u} - \mathbf{u}^\varepsilon||_{1,r} \leqslant C\varepsilon^{\frac{1}{(r-1)(3-r)}} \qquad \text{for } 1 < r \leqslant 2$$

$$\text{and} \quad ||\mathbf{u} - \mathbf{u}^\varepsilon||_{1,r} \leqslant C\varepsilon^{\frac{1}{(r-1)^2}} \qquad \text{for } 2 \leqslant r < \infty.$$

Furthermore,

$$||p - p^\varepsilon||_{0,r'} \leqslant C\varepsilon^{\frac{1}{(3-r)}} \qquad \text{for } 1 < r \leqslant 2$$

$$\text{and} \quad ||p - p^\varepsilon||_{0,r'} \leqslant C\varepsilon^{\frac{1}{(r-1)^2}} \qquad \text{for } 2 \leqslant r < \infty.$$

The above results will be used in the following section for the corresponding fully discrete finite element convergence and error analysis.

4. FINITE ELEMENT APPROXIMATION USING PENALTY METHOD

Let $\mathscr{T}_h$ be a simplicial subdivision of Ω with maximum mesh size

$$h = \max_{K\in\mathscr{T}_h} \operatorname{diam}(K).$$

In this discretization, K denotes a d-simplex, $\operatorname{diam}(K)$ denotes the diameter of K, and ρ_K denotes the radius of the largest closed ball contained in $\overline{K}$. It is assumed

here that $\mathscr{T}_h$ is regular, i.e., there exists a positive constant γ independent of h such that

$$\max_{K \in \mathscr{T}_h} \frac{\text{diam}(K)}{\rho_K} \leqslant \gamma.$$

Let $S_0^h(\Omega)$ be the standard C^0 finite element space in $W_0^{1,r}(\Omega)$ consisting of piecewise polynomials of degree k in Ω. Thus, $\forall v \in S_0^h(\Omega)$ and $\forall K \in \mathscr{T}_h$, $v|_K \in \mathscr{P}_k^d$ where $\mathscr{P}_k^d$ denotes the set of k degree polynomials in d variables. See Ciarlet [23] for further details. The product finite element space $\mathbf{S}_0^h(\Omega) = [S_0^h(\Omega)]^d$ is equipped with the natural product norm. Let Π be the interpolation operator defined by Scott and Zhang, [58] which is associated with $S_0^h(\Omega)$. This operator differs from standard Lagrange interpolation by using local averaging to generate nodal values for functions in Sobolev spaces which may not be pointwise well defined. For interpolation of vectored valued functions, $\Pi \mathbf{u} = (\Pi u_1, \ldots, \Pi u_d)$ for $\mathbf{u} \in \mathbf{W}_0^{1,r}(\Omega)$ is used. An estimate of the interpolation error in Sobolev norms is needed in the following analysis. Let $\mathbf{v} \in \mathbf{W}^{\ell,\alpha}(\Omega)$, where $\ell > \frac{1}{\alpha}$ and $\alpha > 1$. Suppose further that $1 < \beta < +\infty$, $0 \leqslant m \leqslant \ell \leqslant k+1$, and there is a constant σ satisfying

$$0 < \sigma \leqslant \frac{1}{\beta} - \frac{1}{\alpha} + \frac{\ell - m}{d}.$$

Then we have

$$||\mathbf{v} - \Pi\mathbf{v}||_{m,\beta} \leqslant Ch^{\ell - m + d\left(\frac{1}{\beta} - \frac{1}{\alpha}\right)}||\mathbf{v}||_{\ell,\alpha}, \tag{4.1}$$

where C is a positive constant dependent only on Ω. This estimate is a slightly generalized version of Theorem 4.1 of Scott and Zhang [58] where it is proved for the case $\alpha = \beta$ and for a scalar valued v instead of the vector valued $\mathbf{v}$. For completeness, we state their result here: If $\ell > \frac{1}{\beta}$, where $\beta > 1$, then if $0 \leqslant m \leqslant \ell \leqslant k+1$ we have

$$\sum_{K \in \mathscr{T}_h} ||v - \Pi v||_{W^{m,\beta}(K)} \leqslant Ch^{\ell - m}||v||_{\ell,\beta}, \quad \forall v \in W^{\ell,\beta}(\Omega). \tag{4.2}$$

A sketch of the proof for $\alpha \neq \beta$ is given in Wei and Lefton [63]. An interested reader is referred to [63], [58], and [27] for further details. For each $\varepsilon > 0$ and $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$, we consider $\mathbf{u}_h^\varepsilon \in \mathbf{S}_0^h(\Omega)$ to be the unique solution of

$$\min_{\mathbf{u} \in \mathbf{S}_0^h(\Omega)} J_\varepsilon(\mathbf{u}). \tag{4.3}$$

Standard convex analysis can be used to show that such a unique minimizer $\mathbf{u}_h^\varepsilon$ exists and is also the unique solution of the Euler-Lagrange equation

$$\langle A(\mathbf{u}_h^\varepsilon), \mathbf{v} \rangle + \frac{1}{\varepsilon} \langle \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon), \nabla \cdot \mathbf{v} \rangle = \langle \mathbf{f}, \mathbf{v} \rangle \qquad \forall \mathbf{v} \in \mathbf{S}_0^h(\Omega). \tag{4.4}$$

Lemma 4.1. *Let $\mathbf{u}_h^\varepsilon$ be the unique minimizer of (4.3). Then $\exists\,C$ independent of ε and h such that $||\mathbf{u}_h^\varepsilon||_{1,r} \leqslant C$ and $||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r}}$. Furthermore, let $\mathbf{V} = \{\mathbf{v} \in C_0^\infty(\Omega): \nabla\cdot\mathbf{v} = 0\}$, and suppose that $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ satisfies $\forall\varphi \in \mathbf{V}, \langle\mathbf{f},\varphi\rangle = 0$, then $||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r-1}}$*

Proof. Letting $\mathbf{v} = \mathbf{u}_h^\varepsilon$ in (4.4) gives $\langle A(\mathbf{u}_h^\varepsilon),\mathbf{u}_h^\varepsilon\rangle + \frac{1}{\varepsilon}\langle\varphi(\nabla\cdot\mathbf{u}_h^\varepsilon),\nabla\cdot\mathbf{u}_h^\varepsilon\rangle = \langle\mathbf{f},\mathbf{u}_h^\varepsilon\rangle$, which together with (2.4) and (2.5) implies that

$$||\mathbf{u}_h^\varepsilon||_{1,r}^r + \frac{1}{\varepsilon}||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r}^r \leqslant ||\mathbf{f}||_{-1,r'}||\mathbf{u}_h^\varepsilon||_{1,r}.$$

Therefore $||\mathbf{u}_h^\varepsilon||_{1,r}^{r-1} \leqslant C||\mathbf{f}||_{-1,r'}$, $||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r}^r \leqslant C\varepsilon$ and $\nabla\cdot\mathbf{u}_h^\varepsilon \to 0$ in $L^r(\Omega)$ as $\varepsilon \to 0$. Furthermore, if $\forall\varphi \in \mathbf{V}, \langle\mathbf{f},\varphi\rangle = 0$, then by De Rham's Theorem, see Theorem 2.8 of Amrouche and Girault [6], there exists a $q \in L^{r'}(\Omega)$ such that $\mathbf{f} = \nabla q$. Therefore,

$$|\langle\mathbf{f},\mathbf{u}_h^\varepsilon\rangle| = |\langle\nabla q,\mathbf{u}_h^\varepsilon\rangle| = |\langle q,\nabla\cdot\mathbf{u}_h^\varepsilon\rangle| \leqslant ||q||_{0,r'}||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r},$$

and

$$||\mathbf{u}_h^\varepsilon||_{1,r}^r + \frac{1}{\varepsilon}||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r}^r \leqslant ||q||_{0,r'}||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r},$$

which gives $||\nabla\cdot\mathbf{u}_h^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r-1}}$. □

The following Theorem of Adolfsson and Jerison, see [2] and [3], is needed to prove Theorem 4.2 which is the main result of this paper:

Theorem 4.1. *Suppose Ω is a bounded convex domain in $\mathbb{R}^d, d \geqslant 2$ and $f \in L^r(\Omega)$, $1 < r \leqslant 2$ and $\int_\Omega f d\mathbf{x} = 0$. Then there exists a solution w in $W^{2,r}(\Omega)$ solving*

$$\begin{cases} \Delta w = f \text{ in } \Omega, \\ \frac{\partial w}{\partial n} = 0 \text{ on } \partial\Omega, \end{cases}$$

where $\frac{\partial w}{\partial n} = \nabla w \cdot n$ denotes the outer normal derivative of w. Moreover,

$$\int_\Omega |\nabla^2 w|^r d\mathbf{x} \leqslant \int_\Omega |f|^r d\mathbf{x},$$

where ∇^2 denotes the second order derivatives and the constant C depends only on the Lipschitz constant of the domain.

Remark 4.1. Theorem 4.1 does not hold for $2 < r < \infty$ in bounded convex domains. However, if the domain Ω is C^2 or better, then the result holds for all $1 < r < \infty$, see, e.g., Theorem 37-I in the book of Miranda [47].

Theorem 4.2. *Suppose that Ω is a bounded convex domain in $\mathbb{R}^d$ and assume* $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$. *For $0 < \varepsilon \leqslant 1$, let $\mathbf{u}$ be the unique minimizer of the functional J over* $\mathbf{W}_0^{1,r}(\Omega)$, *and let $\mathbf{u}_h^\varepsilon$ be the unique minimizer of $J_\varepsilon(\mathbf{u})$ over $\mathbf{S}_0^h(\Omega)$. Then $\forall \mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$*

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(3-r)}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{(3-r)}} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{2}} + \varepsilon^{\frac{1}{2r}},$$

for $1 < r \leqslant 2$, and

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(r-1)}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{(r-1)}} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{r}} + \varepsilon^{\frac{1}{r^2}},$$

for $2 \leqslant r < \infty$. Furthermore, let $\mathbf{V} = \{\mathbf{v} \in C_0^\infty(\Omega) : \nabla \cdot \mathbf{v} = 0\}$, and if $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ satisfies $\forall \varphi \in \mathbf{V}, \langle \mathbf{f}, \varphi \rangle = 0$. Then $\forall \mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{2(3-r)}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{(3-r)}} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{2}} + \varepsilon^{\frac{1}{2(r-1)}} \tag{4.5}$$

for $1 < r \leqslant 2$.

Note: When $r = 2$, (4.5)) reduces to

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{2}} ||\mathbf{u} - \mathbf{v}_h||_{1,r} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{2}} + \varepsilon^{\frac{1}{2}}.$$

Letting $\varepsilon = h$ and using Scott and Zhang interpolation, we obtain

$$||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant C||\mathbf{u}||_{2,\alpha} h^{\frac{1}{2}}.$$

We point out that the corresponding classical result has an h on the right hand side. A similar gap is also observed by Bercovier and Engelman [14] in studying some other non-Newtonian flows.

Proof. Consider the simple identity:

$$\langle A(\mathbf{u}) - A(\mathbf{u}_h^\varepsilon), \mathbf{u} - \mathbf{u}_h^\varepsilon \rangle = \langle A(\mathbf{u}) - A(\mathbf{u}_h^\varepsilon), \mathbf{u} - \mathbf{v}_h \rangle + \langle A(\mathbf{u}) - A(\mathbf{u}_h^\varepsilon), \mathbf{v}_h - \mathbf{u}_h^\varepsilon \rangle \tag{4.6}$$

which holds for any $\mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$. Let $\mathbf{v} = \mathbf{v}_h - \mathbf{u}_h^\varepsilon$, then by Theorem 4.1, there exists a w such that $\Delta w = \nabla \cdot \mathbf{v}$ satisfying $||\nabla w||_{1,r} \leqslant C||\nabla \cdot \mathbf{v}||_{0,r}$, for $1 < r \leqslant 2$. Similarly, there exists a w such that $||\nabla w||_{1,2} \leqslant C||\nabla \cdot \mathbf{v}||_{0,2} \leqslant C||\nabla \cdot \mathbf{v}||_{0,r}$, for $2 \leqslant r < \infty$. In (3.1), replace the test function by $\mathbf{v} - \nabla w$, which is divergence free, to get

$$\langle A(\mathbf{u}), \mathbf{v}_h - \mathbf{u}_h^\varepsilon \rangle = \langle f, \mathbf{v}_h - \mathbf{u}_h^\varepsilon \rangle + \langle A(\mathbf{u}) - f, \nabla w \rangle \tag{4.7}$$

Similarly, by (4.4),

$$\langle A(\mathbf{u}_h^\varepsilon), \mathbf{v}_h - \mathbf{u}_h^\varepsilon \rangle + \frac{1}{\varepsilon} \langle \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon), \nabla \cdot \mathbf{v}_h - \nabla \cdot \mathbf{u}_h^\varepsilon \rangle$$
$$= \langle f, \mathbf{v}_h - \mathbf{u}_h^\varepsilon \rangle \tag{4.8}$$

Take the difference of (4.7) and (4.8). This allows us to rewrite the last term (4.6) to yield

$$\langle A(\mathbf{u}) - A(\mathbf{u}_h^\varepsilon), \mathbf{u} - \mathbf{u}_h^\varepsilon \rangle = \langle A(\mathbf{u}) - A(\mathbf{u}_h^\varepsilon), \mathbf{u} - \mathbf{v}_h \rangle$$
$$+ \frac{1}{\varepsilon} \langle \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon), \nabla \cdot \mathbf{v}_h - \nabla \cdot \mathbf{u}_h^\varepsilon \rangle + \langle A(\mathbf{u}) - f, \nabla w \rangle \tag{4.9}$$

for any $\mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$. Adding $\frac{1}{\varepsilon} \langle \varphi(\nabla \cdot \mathbf{u}) - \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon), \nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon \rangle$ to both sides of (4.9) gives

$$\langle A(\mathbf{u}) - A(\mathbf{u}_h^\varepsilon), \mathbf{u} - \mathbf{u}_h^\varepsilon \rangle + \frac{1}{\varepsilon} \langle \varphi(\nabla \cdot \mathbf{u}) - \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon), \nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon \rangle$$
$$= \langle A(\mathbf{u}) - A(\mathbf{u}_h^\varepsilon), \mathbf{u} - \mathbf{v}_h \rangle + \frac{1}{\varepsilon} \langle \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon), \nabla \cdot \mathbf{v}_h - \nabla \cdot \mathbf{u} \rangle + \langle A(\mathbf{u}) - f, \nabla w \rangle \tag{4.10}$$

for any $\mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$. Using (4.10), (2.1), and (2.4) we get

$$C \left(\frac{||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}^2}{(||\mathbf{u}||_{1,r} + ||\mathbf{u}_h^\varepsilon||_{1,r})^{2-r}} + \frac{1}{\varepsilon} \frac{||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^2}{(||\nabla \cdot \mathbf{u}||_{0,r} + ||\nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r})^{2-r}} \right)$$
$$\leqslant ||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}^{r-1} ||\mathbf{u} - \mathbf{v}_h||_{1,r} + \frac{1}{\varepsilon} ||\nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^{r-1} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r}$$
$$+ \langle A(\mathbf{u}) - f, \nabla w \rangle \tag{4.11}$$

for $1 < r \leqslant 2$. Similarly, using (2.2) and (2.5) we have, for $2 \leqslant r < \infty$

$$C \left(||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}^r + \frac{1}{\varepsilon} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^r \right)$$
$$\leqslant ||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} (||\mathbf{u}||_{1,r} + ||\mathbf{u}_h^\varepsilon||_{1,r})^{r-2} ||\mathbf{u} - \mathbf{v}_h||_{1,r}$$
$$+ \frac{1}{\varepsilon} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r} (||\nabla \cdot \mathbf{u}||_{0,r} + ||\nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r})^{r-2} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r}$$
$$+ \langle A(\mathbf{u}) - f, \nabla w \rangle. \tag{4.12}$$

Recall that the bounds $||\nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r}}$, $||\mathbf{u}^\varepsilon||_{1,r} \leqslant C$, and $||\mathbf{u}_h^\varepsilon||_{1,r} \leqslant C$, where C is a constant independent of h and ε were proved in Theorem 3.1 and Lemma 4.1. Since $\nabla \cdot \mathbf{u} = 0$ we obtain

$$C \left(||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}^2 + \left(\frac{1}{\varepsilon} \right)^{\frac{2}{r}} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^2 \right) \leqslant ||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}^{r-1} ||\mathbf{u} - \mathbf{v}_h||_{1,r}$$
$$+ \frac{1}{\varepsilon} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^{r-1} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r} + \langle A(\mathbf{u}) - f, \nabla w \rangle \tag{4.13}$$

for $1 < r \leqslant 2$, and

$$C \left(||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}^r + \frac{1}{\varepsilon} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^r \right) \leqslant ||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} ||\mathbf{u} - \mathbf{v}_h||_{1,r}$$
$$+ \left(\frac{1}{\varepsilon} \right)^{\frac{2}{r}} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r} + \langle A(\mathbf{u}) - f, \nabla w \rangle \tag{4.14}$$

for $2 < r < \infty$.

Rewrite the first two terms on the right hand side of (4.13) using the inequality $ab \leqslant \frac{1}{P}(\delta a)^P + \frac{1}{P'}(\frac{b}{\delta})^{P'}$, $a \geqslant 0, b \geqslant 0$, where we choose $P = \frac{2}{r-1}$, $P' = \frac{2}{3-r}$, and $\delta = \delta' \varepsilon^{-\frac{1}{r}}$, where δ' is chosen sufficiently small. Note that δ' actually must be bounded above and below by an appropriate function of ε. After some computation we obtain

$$
\begin{aligned}
C\big(||\mathbf{u} - \mathbf{u}_h^{\varepsilon}||_{1,r}^2 + (\frac{1}{\varepsilon})^{\frac{2}{r}} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^{\varepsilon}||_{0,r}^2\big) &\leqslant ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{2}{(3-r)}} \\
+ (\frac{1}{\varepsilon})^{\frac{2}{r(3-r)}} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r}^{\frac{2}{3-r}} &+ \langle A(\mathbf{u}) - f, \nabla w \rangle \\
\leqslant (\frac{1}{\varepsilon})^{\frac{2}{r(3-r)}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{2}{(3-r)}} &+ \langle A(\mathbf{u}) - f, \nabla w \rangle
\end{aligned}
\tag{4.15}
$$

for $1 < r \leqslant 2$. A similar argument with $P = r$, $P' = r'$, and $\delta = \delta' \varepsilon^{\frac{1}{r}}$, where, as before, δ' is chosen sufficiently small and bounded above and below by an appropriate function of ε gives

$$
\begin{aligned}
C\big(||\mathbf{u} - \mathbf{u}_h^{\varepsilon}||_{1,r}^r + \frac{1}{\varepsilon} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h^{\varepsilon}||_{0,r}^r\big) &\leqslant ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{r}{r-1}} \\
+ (\frac{1}{\varepsilon})^{\frac{1}{r-1}} ||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r}^{\frac{r}{r-1}} &+ \langle A(\mathbf{u}) - f, \nabla w \rangle \\
\leqslant (\frac{1}{\varepsilon})^{\frac{1}{r-1}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{r}{r-1}} &+ \langle A(\mathbf{u}) - f, \nabla w \rangle
\end{aligned}
\tag{4.16}
$$

for $2 < r < \infty$. Moreover, we can estimate

$$
\begin{aligned}
|\langle A(\mathbf{u}) - f, \nabla w \rangle| &\leqslant ||A(\mathbf{u}) - f||_{-1,r'} ||\nabla w||_{1,r} \leqslant C ||\nabla \cdot \mathbf{v}||_{0,r} \\
&\leqslant C(||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r} + ||\nabla \cdot \mathbf{u}_h^{\varepsilon}||_{0,r}),
\end{aligned}
$$

for $1 < r \leqslant 2$, and

$$
\begin{aligned}
|\langle A(\mathbf{u}) - f, \nabla w \rangle| &\leqslant ||A(\mathbf{u}) - f||_{-1,r'} ||\nabla w||_{1,2} \\
&\leqslant C\big(||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,2} + ||\nabla \cdot \mathbf{u}_h^{\varepsilon}||_{0,2}\big) \\
&\leqslant C\big(||\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{v}_h||_{0,r} + ||\nabla \cdot \mathbf{u}_h^{\varepsilon}||_{0,r}\big),
\end{aligned}
$$

for $2 \leqslant r < \infty$. Combining the above results with (4.15) and (4.16) yields

$$
C ||\mathbf{u} - \mathbf{u}_h^{\varepsilon}||_{1,r} \leqslant (\frac{1}{\varepsilon})^{\frac{1}{r(3-r)}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{(3-r)}} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{2}} + \varepsilon^{\frac{1}{2r}},
\tag{4.17}
$$

for $1 < r \leqslant 2$, and

$$
C ||\mathbf{u} - \mathbf{u}_h^{\varepsilon}||_{1,r} \leqslant (\frac{1}{\varepsilon})^{\frac{1}{r(r-1)}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{(r-1)}} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{r}} + \varepsilon^{\frac{1}{r^2}},
\tag{4.18}
$$

for $2 \leqslant r < \infty$. Furthermore, if $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ satisfies $\forall \varphi \in \mathbf{V}, \langle \mathbf{f}, \varphi \rangle = 0$, then by Lemma 4.1, $||\nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r-1}}$. Then (4.17) can be improved and give the following estimate

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{2(3-r)}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{(3-r)}} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{2}} + \varepsilon^{\frac{1}{2(r-1)}}, \qquad (4.19)$$

for $1 < r \leqslant 2$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Remark 4.2. If the domain Ω is smooth, then the result of Theorem 4.1 holds even for $2 < r < \infty$ and, in this case, the following holds

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(r-1)^2}} ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{(r-1)}} + ||\mathbf{u} - \mathbf{v}_h||_{1,r}^{\frac{1}{r}} + \varepsilon^{\frac{1}{r(r-1)}}, \qquad (4.20)$$

for $2 \leqslant r < \infty$.

Corollary 4.1. Suppose that Ω is a bounded convex domain in $\mathbb{R}^d$. Let $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ satisfy $\langle \mathbf{f}, \varphi \rangle = 0 \ \forall \varphi \in \mathbf{V}$, where $\mathbf{V} = \{\mathbf{v} \in C_0^\infty(\Omega) : \nabla \cdot \mathbf{v} = 0\}$. For $0 < \varepsilon \leqslant 1$, let $\mathbf{u}$ be the unique minimizer of the functional J over $\mathbf{W}_0^{1,r}(\Omega)$, and let $\mathbf{u}_h^\varepsilon$ be the unique minimizer of $J_\varepsilon(\mathbf{u})$ over $\mathbf{S}_0^h(\Omega)$. Suppose that $\alpha > 1$, $\frac{1}{\alpha} < \ell \leqslant k+1$, and there is a constant σ satisfying $0 < \sigma \leqslant \frac{1}{r} - \frac{1}{\alpha} + \frac{\ell-1}{d}$. If there is a $C > 0$ such that $||\mathbf{u}||_{\mathbf{W}^{\ell,\alpha}(\Omega)} \leqslant C$. Then $\forall \mathbf{v}_h \in \mathbf{S}_0^h(\Omega)$

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{2(3-r)}} h^{[\ell-1+d(\frac{1}{r}-\frac{1}{\alpha})]\frac{1}{(3-r)}} + h^{[\ell-1+d(\frac{1}{r}-\frac{1}{\alpha})]\frac{1}{2}} + \varepsilon^{\frac{1}{2(r-1)}},$$

for $1 < r \leqslant 2$ and if, in addition, Ω is C^2, then

$$C||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r} \leqslant \left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(r-1)^2}} h^{[\ell-1+d(\frac{1}{r}-\frac{1}{\alpha})]\frac{1}{(r-1)}} + h^{[\ell-1+d(\frac{1}{r}-\frac{1}{\alpha})]\frac{1}{r}} + \varepsilon^{\frac{1}{r(r-1)}},$$

for $2 \leqslant r < \infty$.

Proof. Applying Theorem 4.2 with $\mathbf{v}_h = \Pi\mathbf{u}$ and using the estimate in (4.1), the result then follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Theorem 4.3. *Let* $(\mathbf{u}, p)$ *be the solution of* (2.6). *Let* $\mathbf{u}_h^\varepsilon$ *be a solution of* (4.3) *and let* $p_h^\varepsilon = c' - \frac{1}{\varepsilon}\varphi(\nabla \cdot \mathbf{u}_h^\varepsilon)$ *where* $c' = \frac{1}{\varepsilon|\Omega|}\int_\Omega \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon)d\mathbf{x}$. *Suppose that* (1.3) *holds. Then there exists a positive constant C such that*

$$||p - p_h^\varepsilon||_{0,1} \leqslant C(\varepsilon^{\frac{1}{3-r}} + \frac{1}{\varepsilon}||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}^{r-1}),$$

for $1 < r \leqslant 2$, *and*

$$||p - p_h^\varepsilon||_{0,1} \leqslant C(\varepsilon^{\frac{1}{(r-1)^2}} + \frac{1}{\varepsilon^{\frac{2}{r}}}||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,r}),$$

for $2 \leqslant r < \infty$.

Proof. By (2.1), (2.2) and Hölder's inequality, we have

$$\int_\Omega |\varphi(\nabla \cdot \mathbf{u}^\varepsilon) - \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon)| d\mathbf{x} \leqslant C ||\nabla \cdot \mathbf{u}^\varepsilon - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^{r-1}, \qquad (4.21)$$

for $1 < r \leqslant 2$ and

$$\int_\Omega |\varphi(\nabla \cdot \mathbf{u}^\varepsilon) - \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon)| d\mathbf{x} \leqslant$$
$$C(||\nabla \cdot \mathbf{u}^\varepsilon||_{0,r} + ||\nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r})^{r-2} ||\nabla \cdot \mathbf{u}^\varepsilon - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}, \qquad (4.22)$$

for $2 \leqslant r < \infty$. Therefore, using the definitions of p^ε and p_h^ε

$$\int_\Omega |p^\varepsilon - p_h^\varepsilon| d\mathbf{x} \leqslant \frac{C}{\varepsilon} \int_\Omega |\varphi(\nabla \cdot \mathbf{u}^\varepsilon) - \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon)| d\mathbf{x} \leqslant \frac{C}{\varepsilon} ||\nabla \cdot \mathbf{u}^\varepsilon - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}^{r-1}, \quad (4.23)$$

for $1 < r \leqslant 2$. And since $||\nabla \cdot \mathbf{u}^\varepsilon||_{0,r} + ||\nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r} \leqslant C\varepsilon^{\frac{1}{r}}$, we have

$$\int_\Omega |p^\varepsilon - p_h^\varepsilon| d\mathbf{x} \leqslant \frac{C}{\varepsilon} \int_\Omega |\varphi(\nabla \cdot \mathbf{u}^\varepsilon) - \varphi(\nabla \cdot \mathbf{u}_h^\varepsilon)| d\mathbf{x} \leqslant \frac{C}{\varepsilon^{\frac{2}{r}}} ||\nabla \cdot \mathbf{u}^\varepsilon - \nabla \cdot \mathbf{u}_h^\varepsilon||_{0,r}, \quad (4.24)$$

for $2 \leqslant r < \infty$. By (4.23) and (4.24),

$$||p - p_h^\varepsilon||_{0,1} \leqslant ||p - p^\varepsilon||_{0,1} + ||p^\varepsilon - p_h^\varepsilon||_{0,1} \leqslant ||p - p^\varepsilon||_{0,1} + \frac{C}{\varepsilon} ||\mathbf{u}^\varepsilon - \mathbf{u}_h^\varepsilon||_{1,r}^{r-1},$$

for $1 < r \leqslant 2$, and

$$||p - p_h^\varepsilon||_{0,1} \leqslant ||p - p^\varepsilon||_{0,1} + ||p^\varepsilon - p_h^\varepsilon||_{0,1} \leqslant ||p - p^\varepsilon||_{0,1} + \frac{C}{\varepsilon^{\frac{2}{r}}} ||\mathbf{u}^\varepsilon - \mathbf{u}_h^\varepsilon||_{1,r},$$

for $2 \leqslant r < \infty$.

This completes the proof of the theorem since $||\nabla \cdot \mathbf{v}||_{0,r} \leqslant C ||\mathbf{v}||_{1,r} \, \forall \mathbf{v} \in \mathbf{W}_0^{1,r}(\Omega)$, and $||\mathbf{u} - \mathbf{u}^\varepsilon||_{1,r}$ is estimated by Theorem 3.2. $\qquad \square$

Corollary 4.2. Suppose $\alpha > 1$, $\frac{1}{\alpha} < \ell \leqslant k+1$, and there is a constant σ satisfying $0 < \sigma \leqslant \frac{1}{r} - \frac{1}{\alpha} + \frac{\ell-1}{d}$. Suppose there is a $C > 0$ independent of ε and h such that $||\mathbf{u}^\varepsilon||_{\mathbf{W}^{\ell,\alpha}(\Omega)} \leqslant C$. Then

$$||p - p_h^\varepsilon||_{0,1} \leqslant C \left(\varepsilon^{\frac{1}{3-r}} + \varepsilon^{\frac{-r^2+4r-1}{r(3-r)}} h^{\frac{r-1}{3-r}[\ell-1+d(\frac{1}{r}-\frac{1}{\alpha})]} \right)$$

for $1 < r \leqslant 2$ and if, in addition, Ω is C^2, then

$$||p - p_h^\varepsilon||_{0,1} \leqslant C \left(\varepsilon^{\frac{1}{(r-1)^2}} + \varepsilon^{-\frac{2r-1}{r(r-1)}} h^{\frac{1}{r-1}[\ell-1+d(\frac{1}{r}-\frac{1}{\alpha})]} \right)$$

for $2 \leqslant r < \infty$.

Proof. The results follow from Applications of Theorem 4.3, Corollary 4.1, and Theorem 4.2. $\qquad\square$

Remark 4.3. The rates of convergence in the error estimates given by Corollary 4.1 and Corollary 4.2 depend on both ε and h. However, if $\varepsilon > 0$ is fixed, the convergence of $\mathbf{u}_h^\varepsilon$ to $\mathbf{u}^\varepsilon$ as $h \to 0$ may be obtained by choosing $h = \varepsilon^{\frac{1}{\gamma}}$ with $0 < \gamma < \left[\ell - 1 + d(\frac{1}{r} - \frac{1}{\alpha})\right] \frac{\vartheta_2}{\vartheta_1}$. However, this requires a very fine mesh when ε is large and the rate of convergence is not optimal. This is also observed in the numerical experiments. In addition, the boundedness of $||\mathbf{u}||_{\mathbf{W}^{\ell,\alpha}(\Omega)}$ requires potentially higher regularity on the solution. It has been shown, of course, that $||\mathbf{u}||_{\mathbf{W}^{1,r}(\Omega)} \leqslant C$, but that is not sufficient for a general strong error estimate.

When $r = 2$, for the Newtonian flows, we have the following Theorem due to Temam [59]:

Theorem 4.4. *Let $\ell \geqslant 2$ be an integer, α any real number with $1 < \alpha < \infty$ and let Ω be a bounded and open set of $\mathbb{R}^d$ of class C^ℓ. Let*

$$\mathbf{f} \in \mathbf{W}^{\ell-2,\alpha}(\Omega), \psi \in W^{\ell-1,\alpha}(\Omega), and\, \mathbf{g} \in \mathbf{W}^{\ell-\frac{1}{\alpha},\alpha}(\partial\Omega)$$

be given with $\mathbf{g}$ and psi satisfying the compactibility condition

$$\int_{\partial\Omega} \mathbf{g} \cdot \mathbf{n} d\mathbf{x} = \int_\Omega \psi d\mathbf{x}.$$

Then, the nonhomogeneous Stokes problem

$$\begin{aligned}
-k\Delta\mathbf{u} + \nabla p &= \mathbf{f} &&in\ \Omega, \\
\nabla \cdot \mathbf{u} &= \psi &&in\ \Omega, \\
\mathbf{u} &= \mathbf{g} &&on\ \partial\Omega.
\end{aligned} \tag{4.25}$$

has a unique solution $(\mathbf{u}, p)$ in $\mathbf{W}^{\ell,\alpha}(\Omega) \times W^{\ell-1,\alpha}(\Omega)$. In addition, there exists a constant $C > 0$ dependent only on ℓ, α, k, and Ω, such that

$$||\mathbf{u}||_{\ell,\alpha} + ||p||_{\ell-1,\alpha} \leqslant C(||\mathbf{f}||_{\ell-2,\alpha} + ||\psi||_{\ell-1,\alpha} + ||\mathbf{g}||_{\ell-\frac{1}{\alpha},\alpha}).$$

With the uniform bound on $||\mathbf{u}||_{\ell,\alpha}$, one can then use Corollary 4.1 and Corollary 4.2 to give the following error estimates for the Newtonian Stokes problem:

$$\begin{aligned}
||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,2} &\leqslant C\left(\varepsilon + \varepsilon^{-\frac{1}{2}} h^{\ell-1+d(\frac{1}{2}-\frac{1}{\alpha})}\right), \\
||p - p_h^\varepsilon||_{0,1} &\leqslant C(\varepsilon + \varepsilon^{-\frac{3}{2}} h^{\ell-1+d(\frac{1}{2}-\frac{1}{\alpha})}).
\end{aligned} \tag{4.26}$$

In particular, let $\varepsilon = h$ and $\ell = 2$ in (4.26), then we obtain the optimal error estimate in velocity for the linear finite elements:

$$||\mathbf{u} - \mathbf{u}_h^\varepsilon||_{1,2} = O(h^{\frac{1}{2} + d(\frac{1}{2} - \frac{1}{\alpha})}). \tag{4.27}$$

This is comparable to the rate of $O(h^{\frac{1}{2}})$ found in Reddy [53] under similar conditions. The regularity theory presented in [60], [25], and [21] indicate that for appropriate smooth Ω and $\mathbf{f}$, the solution $\mathbf{u}$ may be in $C^{1,\mu}(\bar{\Omega})$ for some $0 < \mu < 1$. Therefore, $||\mathbf{u}||_{1,\alpha} \leqslant C$ with C independent of ε and for any $1 < \alpha < \infty$. Let $\ell = 1$ and $\alpha \to \infty$ in the above error estimates, one can then replace the term $\ell - 1 + d(\frac{1}{r} - \frac{1}{\alpha})$ in (4.26) by $\frac{d}{r}$.

5. NUMERICAL SOLUTIONS

Let $\{N_i\}_{i=1,m}$ be the global basis for the finite element space $\mathbf{S}_0^h(\Omega)$. Then

$$\mathbf{S}_0^h(\Omega) = \{\mathbf{v} | \mathbf{v} = \sum_i^m x_i N_i, \mathbf{x} = (x_1, x_2, ..., x_m) \in \mathbb{R}^m\}.$$

Consider the unconstrained nonlinear program: $\min_{\mathbf{x} \in \mathbb{R}^m} f_\varepsilon(\mathbf{x})$, where $f_\varepsilon(\mathbf{x}) = J_\varepsilon(\mathbf{v})$ and $\mathbf{v} = \sum_i^m x_i N_i \in \mathbf{S}_0^h(\Omega)$. Our purpose is to present some numerical solutions by using the penalty method. For simplicity, we only test the method by using the linear triangular elements on a cavity Stokes flow problem. The following hybrid nonlinear conjugate gradient scheme is used for finding the minimizer of the above nonlinear program:

Set $p_{-1} = 0$, $\beta_0 = 0$, $\mathbf{x}_0 = 0$, and set the convergence tolerance $\delta > 0$.
For $i = 0, 1, ...$
If $||\nabla f(\mathbf{x}_i)|| < \delta$ stop.
If $i > 0$ set
$y_{i-1} = \nabla f(\mathbf{x}_i) - \nabla f(\mathbf{x}_{i-1})$
$\beta_i^{HS} = y_{i-1}^T \nabla f(\mathbf{x}_i) / y_{i-1}^T p_{i-1}$
$\beta_i^{DY} = \nabla f(\mathbf{x}_i)^T \nabla f(\mathbf{x}_i) / y_{i-1}^T p_{i-1}$
$\beta_i = \max\{0, \min\{\beta_i^{HS}, \beta_i^{DY}\}\}$
end if
Set $p_i = -\nabla f(\mathbf{x}_i) + \beta_i p_{i-1}$
Use a line search to determine α_i and set $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i p_i$

In the above algorithm, a back-tracking line search method is used and the step length α_k to be the first element of the sequence: $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, ..., 2^{-i}, ...$ that satisfies a sufficient decrease condition, see, e.g., Nash and Sofer [48] for details. This hybrid nonlinear conjugate gradient method was first proposed by Touati-Ahmed and Storey [61]. It has been tested to perform well for many difficult numerical examples
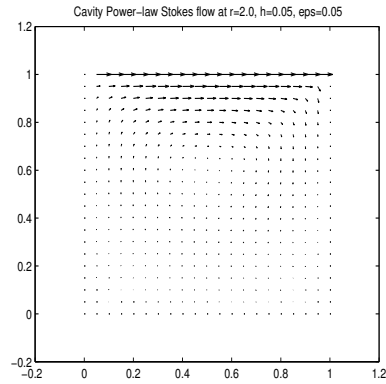
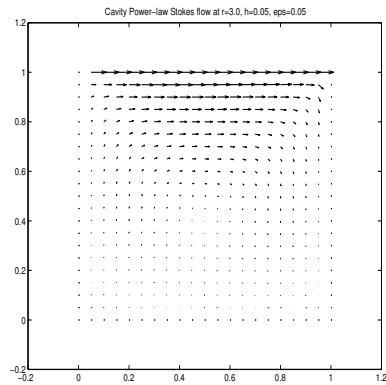**Figure 1.** Newtonian Stokes flow with $r = 2.0$

Figure 2. non-Newtonian Stokes flow with $r = 3.0$

[24] in comparison with the Polak-Ribiére-Polyak method and it does not require the line search scheme to satisfy the strong Wolfe condition. The numerical example is problem (3.2) with $\Omega = (0,1) \times (0,1)$, $k = 1$, $f = 0$, and boundary conditions $\mathbf{u}_\varepsilon = (1,0)$, on $(0,1) \times \{1\}$ and $\mathbf{u}_\varepsilon = (0,0)$, otherwise. Figure 1 is the graphic solution of the problem for the Newtonian Stokes flow with $r = 2.0$. Figure 2 is the graphic solution of the problem for the corresponding power-law non-Newtonian Stokes flow with $r = 3.0$. In both of these numerical experiments $\varepsilon = .05$ and $h = .05$. These numerical experiments support our analysis. Oscillations in the gradient of the objective function is observed. One apparent difference in the two cases is the center of rotation. For the shear thinning flow, the center is lower, which coincides with physical intuition.

REFERENCES

1. ADAMS, R. A. (1978) *Sobolev Spaces*, Academic Press: San Diego.

2. ADOLFSSON, V. & JERISON, D. (1994) L^p−integrability of the second order derivatives for the Neumann problem in convex domains, *Indiana University Mathematics Journal*, **43**, No. 4, 1123–1138.

3. ADOLFSSON, V. & JERISON, D. (1999) Private communications.

4. AGMON, S., DOUGLIS, A. & NIRENBERG, L. (1964) Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II, *Comm. Pure Appl. Math.*, **17**, 35–92.

5. AMROUCHE, C. & GIRAULT, V. (1990) Propriétés fonvtionnelles d'operateurs application au problème de Stokes en dimension quelconque, *Publications du labratoire d'Analye Numérique*, no. 90025, Université Pierre et Marie Curie.

6. AMROUCHE, C. & GIRAULT, V. (1994) Decomposition of vector spaces and application to the Stokes problem in arbitrary dimension, *Czech. Math. J.* **44**, 109-140.

7. BABŮSKA, I. (August 1971) The finite element method with penalty, *Tech. Note BN*-710, The Institute for Fluid Dynamics and Applied Mathematics, University of Maryland.

8. BARANGER, J. & NAJIB, K. (1990) Analyse numerique des ecoulements quasi-Newtoniens dont la viscosite obeit a la loi puissance ou loi de carreau, *Numer. Math.*, **58**, 34–49.

9. BARANGER, J., GEORGET, P. & NAJIB, K. (1987) Error estimates for a mixed finite element method for a non-Newtonian flow, *J. of Non-Newtonian Fluid Mechanics*, **23**, 415–421.

10. BARRETT, J. W. & LIU, W. B. (1993) Finite element approximation of the *p*-Laplacian, *Math. Comp.*, **61**, 523–537.

11. BARRETT, J. W. & LIU, W. B. (1993) Finite element error analysis of a quasi-Newtonian flow obeying the Carreau or power law, *Numer. Math.*, **64**, 433–453.

12. BARRETT, J. W. & LIU, W. B. (1994) Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow, *Numer. Math.*, **68**, 437–456.

13. BERCOVIER, M. (1978) Perturbation of mixed variational problems. Application to mixed finite element methods, *R.A.I.R.O. Analyse Numérique*, **12**, 211–236.

14. BERCOVIER, M. & ENGELMAN, M. (1979) A finite-element for the numerical solution of viscous incompressible flows, *J. Comput. Phys.*, **30**, 181–210.

15. BERCOVIER, M. & ENGELMAN, M. (1980) A finite-element method for incompressible non-Newtonian flows, *J. Comput. Phys.*, **36**, 313–326.

16. BERCOVIER, M. & PIRONNEAU, O. (1979) Error estimates for finite element method solution of the Stokes problem in the primitives variables, *Numer. Math.*, **33**, 211–224.

17. BERNHARDT, E., BERTACCHI, G. & MORONI, A. (1992) Modeling of flow in extruder Dies - fundamentals and applications of the TMconcept-faBest finite element flow analysis, in *Applications of Computer Modeling of Extrusion and Other Continuous Polymer Processes*, Keith T. O'Brien ed., Oxford University Press: Munich.

18. BOGER, D. V., CABELLI, A. & HALMOS, A. L. (1975) The behavior of a power-law fluid flowing through a sudden expansion, *A. I. Ch. E. Journal*, **21**, 540–549.

19. BREZZI, F. (1974) On the existence, uniqueness, and approximation of saddle-point problems arising from Lagrangian multipliers. *R.A.I.R.O. Anal. Numer.*, **8**, 129–151.

20. CÉA, J. (1971) *Optimization - Théory and Algorithms*, Dunod: Paris.

21. CHEN, Y. Z. & DIBENEDETTO, E. (1989) Boundary estimates for solutions of nonlinear de-

generate parabolic systems, *J. reine angew. Math.*, **395**, 102–131.

22. CHOW, S. S. (1989) Finite element error estimates for non-linear elliptic equations of monotone type, *Numerische Mathematik*, **54**, 373–393.

23. CIARLET, P. G. (1978) *The Finite Element Method for Elliptic Problems*, North-Holland.

24. DAI, Y. H., ET AL. (1998) Testing different nonlinear conjugate gradient methods, *Research Report*, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences.

25. DIBENEDETTO, E. (1994) *Degenerate Parabolic Equations*, Springer Verlag: Berlin.

26. DRÁBEK, P., KUFNER, A. & NICOLOSI, F. (1996) *Nonlinear Elliptic Equations, Singular and Degenerate Case*, University of West Bohemia in Pilsen.

27. DUPONT, T. & SCOTT, R. (1990) Polynomial approximation of functions in Sobolev spaces, *Math. Comp.*, **34**, 441–463.

28. ENGLAND, P. & JACKSON, J. (1989) Active deformation of the continents, *Ann. Rev. Earth Planet. Sci.*, **17**, 197–226.

29. ENGLAND, P. & MCKENZIE, D. (1983) A thin viscous sheet model for continental deformation, *Geophys. J. R. Astr. Soc.*, **70**, 295–321.

30. ENGLAND, P. & MCKENZIE, D. (1983) Correction to: A thin viscous sheet model for continental deformation, *Geophys. J. R. Astr. Soc.*, **73**, 523–532.

31. FALK, R. (1976) A finite element method for the stationary Stokes equations using trial functions which do not have to satisfy div $\mathbf{v} = 0$, *Math. Comp.*, **30**, 698–702.

32. FALK, R. (1975) An analysis of the penalty method and extrapolation for the stationery Stokes problem, in *Advanced Computer Methods for Partial Differential Equations*, edited by R. Vichnevetsky, AICA Publication, 66–69.

33. FALK, R. S. & KING, J. T. (1975) A penalty and extrapolation method for the stationery Stokes equation, *SIAM Journal of Numerical Analysis*, **13**, 814-829.

34. FLUENT INC. (1998) *Fidap Theory Manual*.

35. FASTOOK, J. L. (1993) The finite-element method for solving conservation equations in glaciology, *Computational Science and Engineering*, **1**, No. 1, 55–67.

36. FUCIK, S. & KUFNER, A. (1980) *Nonlinear Differential Equations*, Elsevier: Amsterdam.

37. GIRAULT, V. & RAVIART, P. A. (1986) *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer Series in Computational Mathematics **5**, Springer Verlag: New York.

38. GLOWINSKI, R. & MARROCO, A. (1975) Sur l'approximation par éléments finis d'ordre un, et la résolution par penalisation-dualité, d'une clase des problémes de Dirichlet non linéaires, *R.A.I.R.O. Anal. Numer.*, **9**, 41–76.

39. HEBEKER, F. K. (1982) The penalty method applied to the instationary Stokes equation, *Appl. Anal.*, **14**, 137–154.

40. HUGHES, T. J. R., LIU, W. & BROOKS, A. (1979) Finite element analysis of incompressible viscous flows by the penalty function formulation, *J. Comput. Phys.*, **30**, 1–60.

41. IGA, M. & REDDY, J. N. (1988) Penalty finite element analysis of free surface flows of power-law fluids. *The Mathematics of Finite Elements and Applications, VI* (Uxbridge, 1987), 423–433, Academic Press: London, 76–80.

42. LADYZHENSKAYA, O. A. (1969) *The Mathematical Theory of Viscous Incompressible Flow*. Revised 2nd ed., Gordon and Breach: New York.

43. LIU, T. J., LIN, H. M. & HONG, C. N. (1988) Comparison of two numerical methods for the solution of non-Newtonian flow in ducts, *Int. J. Numer. Methods Fluids*, **8**, 845–861.

44. LIU, T. J., WEN, S. & TSOU, J. (1994) Three-dimensional finite element analysis of polymeric fluid flow in an extrusion die. Part I: Entrance effect, *Polymer Engineering Science*, **34**, No. 10, 827–834.

45. LEFTON, L. & WEI, D. (2001) A penalty method for approximations of the stationary power-law Stokes problem, *Electron. J. Diff. Eqns.*, **2001** , No. 37, 1–14.

46. MALKUS, D. S. & HUGHES, T. J. R. (1978) Mixed finite element methods-reduced and selective integration techniques: A unification of concepts, *Comput. Meth. Appl. Mech. Eng.*, **15**, 63–81.

47. MIRANDA, C. (1970) *Partial Differential Equations of Elliptic Type*, Second Revised Edition, Springer-Verlag: New York.

48. NASH, S. G. & SOFER, A. (1996) *Linear and Nonlinear Programming*, McGraw-Hill.

49. ODEN, J. T., KIKUCHI, N. & SONG, Y. J. (1982) Penalty-finite element methods for the analysis of Stokesian flows, *Comp. Methods Appl. Mech. Engr.*, **31**, 207–329.

50. ODEN, J. (1982) Penalty method and reduced integration for the analysis of fluids, *Penalty-Finite Element Methods in Mechanics*, ed. J. Reddy, Amer. Soc. Mech. Eng.: New York, 21–32.

51. ODEN, J. (1978) A theory of penalty methods for finite element approximations of highly non-linear problems in continuum mechanics, *Computers and Struct.*, **8**, 445–449.

52. ODEN, J. & JACQUOTTE, O. P. (1984) Stability of some mixed finite element methods for Stokesian flows, *Comput. Meth. Appl. Mech. Eng.*, **43**, 231–247.

53. REDDY, J. N. (1978) On the accuracy and existence of solutions to primitive variable models of viscous incompressible fluids, *Internat. J. Engrg. Sci.*, **16**, 921–929.

54. REDDY, J. N. (1979) On the finite element method with penalty for incompressible fluid flow problems, *The Mathematics of Finite Elements and Applications III*, J. R. Whiteman ed., Academic Press: London, 277–235.

55. REDDY, J. & PADHYE, V. (1988) A penalty finite element model for axisymmetric flows of non-Newtonian fluids, *Numer. Meth. for Part. Diff. Eqns.*, **4**, 33–56.

56. REDDY, J. N., REDDY M. P. & AKAY, H. U. (1992 & 1988) Penalty finite element analysis of incompressible flows using element by element solution algorithms, *Comput. Meth. Appl. Mech. Eng.*, **100**, 169–205 and *Numer. Meth. for Part. Diff. Eqns.*, **4**, 33–56.

57. SONDER, L. J. & ENGLAND, P. C. (1986) Vertical averages of rheology of the continental lithosphere, *Earth Planet. Sci. Lett.*, **77**, 81–90.

58. SCOTT, L. R. & ZHANG, S. (1990) Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comp.*, **54**, 483–493.

59. TEMAM, R. (1984) *Navier-Stokes Equations*, North-Holland: Amsterdam.

60. TOLKSDORF, P. (1983) Everywhere-regularity for some quasilinear systems with a lack of ellipticity, *Ann. Mat. Pura Appl.*, **134**, 241–266.

61. TOUATI-AHMED, D. & STOREY, C. (1990) Efficient hybrid conjugate gradient techniques, *J. Optimization Theory Appl.*, **64**, 379–397.

62. WEI, D. (2001) Penalty approximations to the stationary power-law Navier-Stokes problem, *Numer. Funct. Anal. Optim.*, **22** (5&6), 1–14.

63. WEI, D. & LEFTON, L. (1999) A priori L^p error estimates for Galerkin approximations to porous medium and fast diffusion equations, *Math. Comp.*, **68**, 971–989.

64. WILKINSON, W. (1960) *Non-Newtonian Fluids*, Pergamon Press: New York.

65. ZEIDLER, E. (1985) *Nonlinear Functional Analysis and its Applications. Volume III: Variational Methods and Optimization*, Springer-Verlag: New York.

66. ZIENKIEWICZ, O. (1974) Constrained variational principles and penalty function methods in finite element analysis, in *Lecture Notes in Mathematics*, ed. G. A. Watson, Springer Verlag: Berlin, 207–214.