Numer. Math. (2010) 115:261–287
DOI 10.1007/s00211-009-0277-8

Numerische Mathematik

Error estimate of the P_1 nonconforming finite element method for the penalized unsteady Navier-Stokes equations

Xiliang Lu · Ping Lin

Received: 15 August 2008 / Revised: 18 July 2009 / Published online: 2 December 2009

Abstract We consider a finite element method for the penalty formulation of the time dependent Navier-Stokes equations. Usually the improper choice of the finite element space will lead that the error estimate (inversely) depends on the penalty parameter ϵ. We use the classical P_1 nonconforming finite element space for the spatial discretization. Optimal ϵ-uniform error estimations for both velocity and pressure are obtained.

Mathematics Subject Classification (2000) 65N30 · 76D05

1 Introduction

We consider the unsteady Navier-Stokes equations

$$\mathbf{u}_t - \nu\Delta\mathbf{u} + (\mathbf{u}\cdot\mathbf{grad})\,\mathbf{u} + \mathbf{grad}\,p = \mathbf{f}, \tag{1.1}$$

$$\mathrm{div}\mathbf{u} = 0, \tag{1.2}$$

$$\mathbf{u}|_{t=0} = \mathbf{u}_0, \quad \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \tag{1.3}$$

where Ω is an open bounded domain in R^n, $n = 2$ or 3. Since the pressure p does not appear in the incompressibility equation this system is difficult to solve numerically. A way to overcome the above difficulty is to add a pressure dependent perturbation

X. Lu (✉)
Johann Radon Institute for Computational and Applied Mathematics (RICAM),
Austrian Academy of Sciences, Altenbergerstrasse 69, 4040 Linz, Austria
e-mail: lu_xiliang@hotmail.com

P. Lin
Division of Mathematics, University of Dundee, 23 Perth Road, Dundee DD1 4HN, Scotland, UK
e-mail: plin@maths.dundee.ac.uk

to the incompressible constaint div$\mathbf{u} = 0$. This leads to a number of methods, which include the penalty method, the artificial compressibility method, the pressure stabilization method, the projection method and the sequential regularization method (cf. [1,6,7,11,15,16,23,25]). The penalty method is the simplest and is fundamental to the analysis among these methods. There is also recent report of success of the method for liquid crystal flow computations (See [18,19]). Instead of solving system (1.1)–(1.3), we solve $(\mathbf{u}^\epsilon, p^\epsilon)$ from the penalized system:

$$\mathbf{u}_t^\epsilon - \nu\Delta\mathbf{u}^\epsilon + B\left(\mathbf{u}^\epsilon, \mathbf{u}^\epsilon\right) + \mathbf{grad}\, p^\epsilon = \mathbf{f}, \tag{1.4}$$

$$\mathrm{div}\mathbf{u}^\epsilon + \epsilon p^\epsilon = 0, \tag{1.5}$$

$$\mathbf{u}^\epsilon|_{t=0} = \mathbf{u}_0, \quad \mathbf{u}^\epsilon|_{\partial\Omega} = \mathbf{0}, \tag{1.6}$$

where $B(\mathbf{u}, \mathbf{v}) = (\mathbf{u}\cdot\mathbf{grad})\mathbf{v} + \frac{1}{2}(\mathrm{div}\mathbf{u})\mathbf{v}$ is the modified nonlinear convection term, to ensure the whole system dissipative (cf. [24]). We can eliminate pressure p^ϵ from system (1.4)–(1.6) to get an equation which only involves the unknown function $\mathbf{u}^\epsilon$. It makes spatial discretization easier since we only need to discretize the velocity space. The penalized system is related to the elastics equation where the material is nearly incompressible, that is the ratio of Lamé constant μ/λ is very small (See [4]). In [23] a temporal discretization of the penalty system is analyzed and its error is estimated. However, for a spatial discretization there will be different difficulties. When we apply finite element method for the spatial discretization of the penalized system (where no pressure variable explicitly appears), the finite element space still can not be chosen arbitrarily. For instance if we choose the conforming P_1 finite element space, the error bound will be inversely proportional to $\sqrt{\epsilon}$. It is known as a locking phenomenon (cf. [4,20]) and exists in other related methods such as the sequential regularization method (See [16,17,20,21]).

The underlying reason for the locking phenomenon of conforming P_1 finite element is as follows. We let $\mathbf{V}_h$ be the conforming P_1 finite element space, M_h be the piecewise constant finite element space. Then the divergence operator is a linear operator $\mathrm{div} : \mathbf{V}_h \to M_h$. It is known that the kernel of the divergence operator in the conforming P_1 element space contains only $\mathbf{0}$ function (cf. [9]). Hence, for any function $\mathbf{u}_h \in \mathbf{V}_h$, we have

$$\|\mathbf{u}_h\| \le \alpha_h \|\mathrm{div}\mathbf{u}_h\|, \tag{1.7}$$

where α_h is a constant which may depend on h, but does not depend on ϵ. Denote the solution to the continuous penalized Navier-Stokes equation and to the finite element solution by $\mathbf{u}^\epsilon$ and $\mathbf{u}_h^\epsilon$, respectively. From [23], it is known

$$\|\mathbf{u}^\epsilon\| = O(1), \quad \|\mathrm{div}\mathbf{u}^\epsilon\| = O(\epsilon).$$

Let $\mathbf{u}_h^\epsilon \in \mathbf{V}_h$ be an approximation of $\mathbf{u}^\epsilon$. Then $\mathbf{u}^\epsilon$ and div$\mathbf{u}^\epsilon$ cannot be approximated well (i.e. error bound independent of the parameter ϵ) by $\mathbf{u}_h^\epsilon$ and div$\mathbf{u}_h^\epsilon$ simultaneously because $\mathbf{u}_h^\epsilon$ has to satisfy (1.7). To avoid the locking, we need to consider some other type of finite element spaces. In [12,13], the conforming mixed finite element approximation for $(\mathbf{u}^\epsilon, p^\epsilon)$ is used, where the finite element space $\mathbf{X}_h$ for the velocity and

M_h for the pressure are needed explicitly. For the proper choice of the finite element spaces $\mathbf{X}_h$ and M_h (for instance Taylor-hood element), optimal ϵ uniform error estimates can be obtained, see [12]. For the improper choice of $\mathbf{X}_h$ and M_h, the error bound will be inversely proportional to $\sqrt{\epsilon}$, see [13] in the case of the steady Navier-Stokes equations.

In this article, we will use a standard nonconforming P_1 element for the velocity (cf. [9,5]) in the penalized system. The pressure does not appear explicitly in the system but it can be recovered by $-\frac{1}{\epsilon}\text{div}\mathbf{u}$. Since the basis function of the nonconforming P_1 finite element space is continuous at only the barycenter of inter-element boundaries (for instance the middle point of the edge of two neighboring elements in 2D). It gives us more freedom to choose a proper approximation for a divergence free function.

We will consider the penalized Navier-Stokes equations by eliminating the pressure variable. For simplicity, we let $\nu = 1$, and denote $\mathbf{u}^\epsilon$ by $\mathbf{u}$,

$$\mathbf{u}_t - \frac{1}{\epsilon}\mathbf{grad}\,\text{div}\mathbf{u} - \Delta\mathbf{u} + B(\mathbf{u},\mathbf{u}) = \mathbf{f}, \tag{1.8}$$

$$\mathbf{u}|_{t=0} = \mathbf{u}_0, \quad \mathbf{u}|_{\partial\Omega} = \mathbf{0}. \tag{1.9}$$

Its variational formulation can be written as

$$(\mathbf{u}_t,\mathbf{v}) + \frac{1}{\epsilon}(\text{div}\mathbf{u},\text{div}\mathbf{v}) + (\mathbf{gradu},\mathbf{gradv}) + b(\mathbf{u},\mathbf{u},\mathbf{v}) = (\mathbf{f},\mathbf{v}), \quad \forall\mathbf{v}\in\mathbf{H}_0^1,$$

where $\mathbf{u}(0) = \mathbf{u}_0$, $b(\mathbf{u},\mathbf{v},\mathbf{w}) = \langle B(\mathbf{u},\mathbf{v}),\mathbf{w}\rangle$ and $\langle\cdot,\cdot\rangle$ represents the duality pair of $\mathbf{H}^{-1}$ and $\mathbf{H}_0^1$. The finite element approximation is more or less to choose a finite element space $\mathbf{V}_h$ which approximates $\mathbf{H}_0^1$. Since we consider the nonconforming case, the derivative will be piecewisely defined. The finite element approximation $\mathbf{u}_h \in \mathbf{V}_h$ satisfies following variational form

$$(\mathbf{u}_{h,t},\mathbf{v}) + a_h(\mathbf{u}_h,\mathbf{v}) + b_h(\mathbf{u}_h,\mathbf{u}_h,\mathbf{v}) = (\mathbf{f},\mathbf{v}), \quad \forall\mathbf{v}\in\mathbf{V}_h,$$

$$\mathbf{u}_h(0) = P_h\mathbf{u}_0.$$

The meaning of a_h, b_h, P_h and V_h will be given in the later section. This paper is to provide an optimal error estimation for both velocity $\mathbf{u}$ and pressure $p = -\frac{1}{\epsilon}\text{div}\mathbf{u}$. The result stands as

$$\|\mathbf{u}(t,\cdot) - \mathbf{u}_h(t,\cdot)\| \le h^2C, \ \forall t\in[0,T], \tag{1.10}$$

$$\|\text{div}_h(\mathbf{u}(t,\cdot) - \mathbf{u}_h(t,\cdot))\| \le \epsilon C(h + t^{-1}h^2), \ \forall t\in[0,T]. \tag{1.11}$$

The estimation for divergence operator has a singularity at the initial time. But the singularity will quickly disappear after $O(h)$ time. This is reasonable since our initial velocity approximation is chosen as an $\mathbf{L}^2$ projection $P_h\mathbf{u}_0$, and it is possible that $\text{div}_h P_h\mathbf{u}_0$ is far from zero even if $\mathbf{u}_0$ is divergence free.

The rest of the paper will be organized as follows. We will introduce notations and a priori estimation for the solution of penalized Navier-Stokes equations in Sect. 2. Some preliminary results for the nonconforming P_1 finite element space will be described in Sect. 3. The main result is given in Sect. 4, and contains the error estimate for the velocity and how we recover the ϵ-uniform error estimate for pressure.

2 Preliminary

2.1 Notation

Let Ω be a convex polygon ($n = 2$) or polyhedron ($n = 3$). We will use f, g, etc. to describe scalar functions, and $\mathbf{f}$, $\mathbf{g}$, etc. to present vector value functions. $L^p(\Omega)$ $(1 \leq p < \infty)$ denotes the space of pth power summable functions in Ω which equips norm

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f|^p dx\right)^{1/p}.$$

If Ω is fixed, we denote it as L^p for simplicity. The space $W^{m,p}$ is a collection of functions which is pth power summable up to mth (weak) derivatives. It can also be viewed as the completion of $C^\infty(\Omega)$ under norm:

$$\|f\|_{W^{m,p}} = \left(\sum_{0\leq\alpha\leq m} \|D^\alpha f\|^p\right)^{1/p}.$$

For any function in $W^{m,p}$, we also define a semi-norm

$$|f|_{m,p} = \left(\sum_{\alpha=m} \|D^\alpha f\|^p\right)^{1/p}.$$

When $p = 2$, $W^{m,2}$ is a Hilbert space. The inner product in L^2 is denoted as $(\cdot, \cdot)$ and its norm as $\|\cdot\|$. We use L^2_0 to indicate the function which is square summable with zero mean. We also replace $W^{m,2}$ by H^m for simplicity. In particular, H^1_0 is the completion of C^∞_0 in H^1, and H^{-1} is considered as the dual space of H^1_0. A vector value function $\mathbf{f} \in \mathbf{W}^{m,p}$ is equivalent to each of its component in the space $W^{m,p}$. When $p = \infty$, we use L^∞ to present the essentially bounded function space with norm

$$\|f\|_{L^\infty} = \text{ess sup}\,|f(x)|.$$

We will repeatedly use the following inequalities and their discrete version. We denote the genetic constant as c which does not depends on the choice of u.

- Poincaré inequality: For any $u \in H^1_0$ or $u \in H^1$ with zero mean, we have

$$\|u\| \leq \frac{d}{\pi}\|\mathbf{grad} u\|, \tag{2.1}$$

  where Ω is a convex domain with diameter d (cf. [22]).

- Hölder inequality:

$$\int_\Omega |fgh| \le \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r} \tag{2.2}$$

  where $p, q, r > 1$ and $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$.
- Sobolev inequality:

$$\|u\|_{L^4} \le c\|u\|^{1/2}\|u\|_1^{1/2},$$
$$\|u\|_{L^\infty} \le c\|u\|^{1/2}\|u\|_2^{1/2},$$

  where $\Omega \in R^2$;

$$\|u\|_{L^6} \le c\|u\|_1,$$
$$\|u\|_{L^\infty} \le c\|u\|_1^{1/2}\|u\|_2^{1/2},$$

  where $\Omega \in R^3$.

2.2 A priori estimates to the penalized Navier-Stokes equations

Define $A_\epsilon \mathbf{u} = -\frac{1}{\epsilon}\mathbf{grad}\,\mathrm{div}\mathbf{u} - \Delta\mathbf{u}$, then we rewrite penalized Navier-Stokes equations as

$$\mathbf{u}_t + A_\epsilon\mathbf{u} + B(\mathbf{u},\mathbf{u}) = \mathbf{f}, \tag{2.3}$$
$$\mathbf{u}|_{t=0} = \mathbf{u}_0, \quad \mathbf{u}|_{\partial\Omega} = \mathbf{0}. \tag{2.4}$$

Lemma 2.1 *For operator A_ϵ, there exists a constant $c_1(\Omega)$, such that if $\epsilon c_1 \le 1$,*

$$\|\mathbf{v}\|_2^2 + \frac{1}{\epsilon^2}\|\mathrm{div}\mathbf{v}\|_1^2 \le c_1\|A_\epsilon\mathbf{v}\|^2, \quad \forall \mathbf{v} \in \mathbf{H}^2 \cap \mathbf{H}_0^1.$$

Proof See Lemma 1.1 of [3]. □

We have some conditions on the initial velocity and external force:

$$\mathrm{div}\mathbf{u}_0 = 0, \quad \|\mathbf{u}_0\|_1 + \sup_{0<t<T}\|\mathbf{f}(\cdot,t)\| \le N_1, \tag{2.5}$$
$$\|\mathbf{u}_0\|_2 + \sup_{0<t<T}\|\mathbf{f}_t(\cdot,t)\| \le N_2, \tag{2.6}$$
$$\sup_{0<t<T} t\|\mathbf{f}_{tt}(\cdot,t)\| \le N_3. \tag{2.7}$$

Lemma 2.2 *We always assume ϵ is small enough, i.e. $\epsilon c_1 \leq 1$, where c_1 is as in Lemma 2.1. If assumption (2.5) holds, $\Omega \in R^3$ (with constant N_1 being small) or $\Omega \in R^2$, then we have*

$$\sup_{0<t<T}\left(\|\mathbf{u}\|_1+\frac{1}{\sqrt{\epsilon}}\|\mathrm{div}\mathbf{u}\|\right)+\left[\int_0^T\left(\|\mathbf{u}\|_2^2+\|\mathbf{u}_t\|^2+\frac{1}{\epsilon^2}\|\mathrm{div}\mathbf{u}\|_1^2\right)dt\right]^{1/2}\leq M_1.$$

If assumptions (2.5) and (2.6) hold, then

$$\sup_{0<t<T}\left(\|\mathbf{u}\|_2+\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1+\|\mathbf{u}_t\|\right)\leq M_2,$$
$$\sup_{0<t<T}\left(\sqrt{t}\|\mathbf{u}_t\|_1+\frac{\sqrt{t}}{\sqrt{\epsilon}}\|\mathrm{div}\mathbf{u}_t\|\right)$$
$$+\left[\int_0^T t\left(\|\mathbf{u}_t\|_2^2+\|\mathbf{u}_{tt}\|^2+\frac{1}{\epsilon^2}\|\mathrm{div}\mathbf{u}_t\|_1^2\right)dt\right]^{1/2}\leq M_3.$$

If assumptions (2.5)–(2.7) hold, then

$$\sup_{0<t<T} t\left(\|\mathbf{u}_t\|_2+\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}_t\|_1+\|\mathbf{u}_{tt}\|\right)\leq M_4,$$
$$\sup_{0<t<T}\left(t^{3/2}\|\mathbf{u}_t\|_1+\frac{\sqrt{t^3}}{\sqrt{\epsilon}}\|\mathrm{div}\mathbf{u}_t\|\right)$$
$$+\left[\int_0^T t^3\left(\|\mathbf{u}_{tt}\|_2^2+\|\mathbf{u}_{ttt}\|^2+\frac{1}{\epsilon^2}\|\mathrm{div}\mathbf{u}_{tt}\|_1^2\right)dt\right]^{1/2}\leq M_5.$$

Proof We will only sketch the proof, while the details are quite similar to [3,14]. Multiplying $\mathbf{u}$ at both sides of Eq. (2.3) and integrating from 0 to T, we have

$$\|\mathbf{u}\|^2+\int_0^T\left(\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|^2+\|\mathbf{u}\|_1^2\right)dt\leq c\left(\|\mathbf{u}_0\|^2+\int_0^T\|\mathbf{f}\|_{-1}^2dt\right). \tag{2.8}$$

Then we multiply $A_\epsilon\mathbf{u}$ to obtain

$$\frac{d}{dt}\left(\|\mathbf{grad}\mathbf{u}\|^2+\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|^2\right)+\|A_\epsilon\mathbf{u}\|^2\leq\begin{cases}c\left(\|\mathbf{grad}\mathbf{u}\|^6+\|\mathbf{f}\|^2\right) & n=3,\\ c\left(\|\mathbf{u}\|^2\|\mathbf{grad}\mathbf{u}\|^4+\|\mathbf{f}\|^2\right) & n=2.\end{cases} \tag{2.9}$$

Hence if $\Omega \in R^3$ (with constant N_1 being small) or $\Omega \in R^2$, the solution $\mathbf{u}$ satisfies $\sup_{0<t<T} \|\mathbf{grad}\,\mathbf{u}(\cdot,t)\| < \infty$. Then we multiply $\mathbf{u}_t$ at both sides of Eq. (2.3), integrate inequality (2.9) and obtain

$$\|\mathbf{u}\|_1^2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|^2 + \int_0^T \left(\frac{1}{\epsilon^2}\|\mathrm{div}\mathbf{u}\|_1^2 + \|\mathbf{u}\|_2^2 + \|\mathbf{u}_t\|^2\right) dt \le M_1. \quad (2.10)$$

To see other estimates, we need to differentiate equation (2.3),

$$\mathbf{u}_{tt} + A_\epsilon \mathbf{u}_t + B(\mathbf{u}, \mathbf{u}_t) + B(\mathbf{u}_t, \mathbf{u}) = \mathbf{f}_t. \quad (2.11)$$

The formula of $\mathbf{u}_t(0)$ can be obtained by choosing $t = 0$ at Eq. (2.3), and $\|\mathbf{u}_t(0)\|$ is bounded when condition (2.6) is satisfied. Multiplying $\mathbf{u}_t$ at both sides of Eq. (2.11), we find

$$\|\mathbf{u}_t\|^2 + \int_0^T \left(\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}_t\|^2 + \|\mathbf{u}_t\|_1^2\right) dt < \infty. \quad (2.12)$$

We then multiply Eq. (2.3) by $A_\epsilon \mathbf{u}$ and conclude from the above inequality that

$$\sup_{0<t<T} \left(\|\mathbf{u}\|_2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1 + \|\mathbf{u}_t\|\right) \le M_2.$$

Multiplying Eq. (2.11) by $tA_\epsilon \mathbf{u}_t$ and integrating with respect to t, we obtain the third estimation. Finally, taking the time derivative of Eq. (2.11) and applying the similar argument we can obtain the last two inequalities. □

Remark 2.1 Let us explain all the constants used in this article. We use c, c_i to represent a genetic constant which only depend on Ω, ν, and use C, M, M_i to represent a constant which may further depend on N_i. It should be noted that any constant here should NOT depend on the choice of the penalty parameter ϵ and the mesh size h.

3 P_1 nonconforming finite element approximation

3.1 Finite element space

Since Ω is a convex polygon or polyhedron, we consider a quasi-uniform partition (cf. [5,10]) $\mathcal{T}_h$ (triangles in R^2 and tetrahedron in R^3), where h represents the maximal diameter of all elements in the partition. Let our finite element space

$$\begin{aligned}\mathbf{V}_h = \{\mathbf{v} : \mathbf{v} \in \mathbf{L}^2(\Omega),\ \mathbf{v}|_T \text{ is linear for all } T \in \mathcal{T}_h,\\ \mathbf{v} \text{ is continuous at the barycenter of inter-element boundaries and}\\ \mathbf{v} = 0 \text{ at the barycenter of edges along } \partial\Omega\},\\ L_h = \{z : z \in L_0^2(\Omega),\ z|_T = \text{constant for all } T \in \mathcal{T}_h\}.\end{aligned}$$

For $\mathbf{v} \in \mathbf{V}_h$, we define $\mathbf{grad}_h\mathbf{v}$ and $\mathrm{div}_h\mathbf{v}$ by

$$(\mathbf{grad}_h\mathbf{v})|_T = \mathbf{grad}(\mathbf{v}|_T), \quad (\mathrm{div}_h\mathbf{v})|_T = \mathrm{div}(\mathbf{v}|_T), \quad \forall T \in \mathcal{T}_h.$$

It is clear that $\mathbf{V}_h$ is not a subspace of $\mathbf{H}_0^1$. The operators $\mathbf{grad}_h$, div_h can be extended to $\mathbf{H}_0^1$ by elementwise calculation, and they are equal to $\mathbf{grad}$, div. Define space $\mathbf{H}$ as the algebraic sum $\mathbf{V}_h \oplus \mathbf{H}_0^1$, then for simplicity, we define a new norm on $\mathbf{H}$ as

$$\|\mathbf{u}\|_{1,h} = \|\mathbf{grad}_h\mathbf{u}\|.$$

By simple calculation, one can verify the operator div_h is from $\mathbf{V}_h$ onto L_h. Moreover, one can show that $(\mathbf{V}_h, L_h)$ satisfies a discrete inf-sup condition:

$$\sup_{\mathbf{v}\in\mathbf{V}_h,\mathbf{v}\neq\mathbf{0}} \frac{|(\mathrm{div}_h\mathbf{v}, q)|}{\|\mathbf{grad}_h\mathbf{v}\|} \geq c\|q\|, \quad \forall q \in L_h. \tag{3.1}$$

This result can be found in [9]. We will use it to recover the accuracy of pressure. We also define another operator $\Delta_h : \mathbf{V}_h \to \mathbf{V}_h$, a discrete Laplace operator by

$$(\Delta_h\mathbf{u}, \mathbf{v}) = -(\mathbf{grad}_h\mathbf{u}, \mathbf{grad}_h\mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h.$$

The finite element space $\mathbf{V}_h$ has inverse estimation (cf. [9])

$$\|\mathbf{grad}_h\mathbf{v}\| \leq ch^{-1}\|\mathbf{v}\|, \quad \forall \mathbf{v} \in \mathbf{V}_h. \tag{3.2}$$

We will repeatedly use bilinear function $(\mathbf{grad}_h\mathbf{u}, \mathbf{grad}_h\mathbf{v}) + \frac{1}{\epsilon}(\mathrm{div}_h\mathbf{u}, \mathrm{div}_h\mathbf{v})$. To simplify notations, let $a_h(\mathbf{u}, \mathbf{v}) = (\mathbf{grad}_h\mathbf{u}, \mathbf{grad}_h\mathbf{v}) + \frac{1}{\epsilon}(\mathrm{div}_h\mathbf{u}, \mathrm{div}_h\mathbf{v})$. Then $a_h(\cdot, \cdot)$ is a symmetric positive definite bilinear form on $\mathbf{H}$. According to [4] we can define the norm induced by a_h as

$$\|\mathbf{u}\|_{a,h} = a_h(\mathbf{u}, \mathbf{u})^{1/2}.$$

Clearly $\|\mathbf{u}\|_{1,h} \leq \|\mathbf{u}\|_{a,h}$ for all $\mathbf{u} \in \mathbf{H}$.

3.2 Various approximate operators

We first define a few extra terms arising from integration by parts due to the nonconforming finite element:

$$X_1(\mathbf{u}, \mathbf{v}) = \int_\Omega \mathbf{grad}\mathbf{u} : \mathbf{grad}_h\mathbf{v}dx + \int_\Omega \Delta\mathbf{u} \cdot \mathbf{v}dx, \tag{3.3}$$

$$X_2(u, \mathbf{v}) = \int_\Omega u \operatorname{div}_h \mathbf{v} dx + \int_\Omega \mathbf{grad} u \cdot \mathbf{v} dx, \tag{3.4}$$

$$X_3(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \frac{1}{2} \sum_{T \in \mathcal{T}_h} \int_{\partial T} (\mathbf{u} \cdot \mathbf{n})(\mathbf{v} \cdot \mathbf{w}) ds, \tag{3.5}$$

where $\mathbf{u} \in \mathbf{H}^2 \cap \mathbf{H}_0^1$, $\mathbf{v}, \mathbf{w} \in \mathbf{H}$.

By using Bramble-Hilbert Lemma (cf. [2]) and homogeneity arguments, we can prove following inequalities (cf. [4,5,14]).

$$|X_1(\mathbf{u}, \mathbf{v})| \leq ch \|\mathbf{u}\|_2 \|\mathbf{grad}_h \mathbf{v}\|, \quad \forall \mathbf{u} \in \mathbf{H}^2 \cap \mathbf{H}_0^1, \mathbf{v} \in \mathbf{H}, \tag{3.6}$$

$$|X_2(u, \mathbf{v})| \leq ch \|u\|_1 \|\mathbf{grad}_h \mathbf{v}\|, \quad \forall u \in H^1 \cap L_0^2, \mathbf{v} \in \mathbf{H}, \tag{3.7}$$

$$|X_3(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq ch \|\mathbf{u}\|_2 \|\mathbf{grad}_h \mathbf{v}\| \|\mathbf{grad}_h \mathbf{w}\|, \quad \forall \mathbf{u} \in \mathbf{H}^2 \cap \mathbf{H}_0^1, \mathbf{v}, \mathbf{w} \in \mathbf{H}. \tag{3.8}$$

The first approximate operator is the interpolation operator $\imath_h : \mathbf{H}^2 \cap \mathbf{H}_0^1 \to \mathbf{V}_h$. It is defined by

$$\imath_h \mathbf{u}(m_e) := \frac{1}{|e|} \int_e \mathbf{u} ds, \tag{3.9}$$

where m_e is the barycenter of edge e. Using the definition of $\mathbf{V}_h$, such a interpolation function is a well-defined piecewise linear polynomial satisfying (3.9) (See [9,4] for more discussion on constructing the interpolation function). Then we have

Lemma 3.1

$$\operatorname{div}(\imath_h \mathbf{u})|_T = \frac{1}{|T|} \int_T \operatorname{div} \mathbf{u} dx, \quad \forall T \in \mathcal{T}_h, \tag{3.10}$$

$$\|\mathbf{u} - \imath_h \mathbf{u}\| + h \|\mathbf{u} - \imath_h \mathbf{u}\|_{1,h} \leq ch^2 \|\mathbf{u}\|_2, \tag{3.11}$$

$$\|\operatorname{div}_h (\mathbf{u} - \imath_h \mathbf{u})\| \leq ch \|\operatorname{div} \mathbf{u}\|_1. \tag{3.12}$$

Proof Equation (3.10) can be verified by definition. Inequality (3.11) is a standard result (see [9]). We will prove (3.12) below. We choose $\mathbf{u}_1 \in \mathbf{H}^2 \cap \mathbf{H}_0^1$ such that (cf. [10])

$$\operatorname{div} \mathbf{u}_1 = \operatorname{div} \mathbf{u}, \quad \|\mathbf{u}_1\|_2 \leq c \|\operatorname{div} \mathbf{u}\|_1. \tag{3.13}$$

From the definition of interpolation operator (3.9) and (3.10), we have

$$(\operatorname{div}_h \imath_h \mathbf{u}_1)|_T = \frac{1}{|T|} \int_T \operatorname{div} \mathbf{u}_1 dx = \frac{1}{|T|} \int_T \operatorname{div} \mathbf{u} dx = (\operatorname{div}_h \imath_h \mathbf{u})|_T,$$

hence

$$\|\mathrm{div}_h(\mathbf{u}-\imath_h\mathbf{u})\| = \|\mathrm{div}_h(\mathbf{u}_1-\imath_h\mathbf{u}_1)\| \le \|\mathbf{u}_1-\imath_h\mathbf{u}_1\|_{1,h} \le ch\|\mathbf{u}_1\|_2 \le ch\|\mathrm{div}\mathbf{u}\|_1.$$

□

The next operator is the a_h orthogonal projection $R_h : \mathbf{H} \to \mathbf{V}_h$ such that

$$a_h(R_h\mathbf{u},\mathbf{v}) = a_h(\mathbf{u},\mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h. \tag{3.14}$$

Evidently this operator is well defined. We will prove that it also satisfies the following Lemma (cf. [4]).

Lemma 3.2 *For all* $\mathbf{u} \in \mathbf{H}^2 \cap \mathbf{H}_0^1$, *we have*

$$\|\mathbf{u}-R_h\mathbf{u}\|_{a,h}^2 \le ch^2\left(\|\mathbf{u}\|_2^2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1^2\right) \tag{3.15}$$

$$\|\mathbf{u}-R_h\mathbf{u}\| + h\|\mathbf{u}-R_h\mathbf{u}\|_{1,h} \le ch^2\left(\|\mathbf{u}\|_2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1\right). \tag{3.16}$$

Proof We notice that $\mathbf{H}$ is a Hilbert space under norm $\|\cdot\|_{a,h}$ and R_h is the a_h orthogonal projection, we have

$$\begin{aligned}\|\mathbf{u}-R_h\mathbf{u}\|_{a,h}^2 &= \inf_{\mathbf{v}\in\mathbf{V}_h}\|\mathbf{u}-\mathbf{v}\|_{a,h}^2 \le \|\mathbf{u}-\imath_h\mathbf{u}\|_{a,h}^2 \\ &= \|\mathbf{u}-\imath_h\mathbf{u}\|_{1,h}^2 + \frac{1}{\epsilon}\|\mathrm{div}_h\,(\mathbf{u}-\imath_h\mathbf{u})\,\|^2 \\ &\le ch^2\left(\|\mathbf{u}\|_2^2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1^2\right).\end{aligned}$$

To obtain the second inequality, we need a duality argument. Let $\mathbf{w}$ solve the following elliptic equation with the homogenous Dirichlet boundary condition:

$$-\Delta\mathbf{w} - \frac{1}{\epsilon}\,\mathbf{grad}\,\mathrm{div}\mathbf{w} = \mathbf{u} - R_h\mathbf{u}. \tag{3.17}$$

Lemma 2.1 gives $\|\mathbf{w}\|_2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{w}\|_1 \le c\|\mathbf{u}-R_h\mathbf{u}\|$. Multiplying $\mathbf{u}-R_h\mathbf{u}$ at both sides of Eq. (3.17) yields

$$\begin{aligned}\|\mathbf{u}-R_h\mathbf{u}\|^2 &= -(\mathbf{u}-R_h\mathbf{u},\Delta\mathbf{w}) - \frac{1}{\epsilon}(\mathbf{u}-R_h\mathbf{u},\mathbf{grad}\,\mathrm{div}\mathbf{w}) \\ &= a_h(\mathbf{u}-R_h\mathbf{u},\mathbf{w}) - X_1\,(\mathbf{w},\mathbf{u}-R_h\mathbf{u}) - \frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{w},\mathbf{u}-R_h\mathbf{u}) \\ &= a_h(\mathbf{u}-R_h\mathbf{u},\mathbf{w}-\imath_h\mathbf{w}) - X_1(\mathbf{w},\mathbf{u}-R_h\mathbf{u}) - \frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{w},\mathbf{u}-R_h\mathbf{u}) \\ &\le \|\mathbf{u}-R_h\mathbf{u}\|_{a,h}\|\mathbf{w}-\imath_h\mathbf{w}\|_{a,h} + ch\left(\|\mathbf{w}\|_2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{w}\|_1\right)\|\mathbf{u}-R_h\mathbf{u}\|_{1,h} \\ &\le ch\|\mathbf{u}-R_h\mathbf{u}\|_{a,h}\|\mathbf{u}-R_h\mathbf{u}\|.\end{aligned}$$

Eliminating $\|\mathbf{u}-R_h\mathbf{u}\|$ from both sides yields the last inequality. □

Another approximate operator is the $\mathbf{L}^2$ projection operator $P_h : \mathbf{L}^2 \to \mathbf{V}_h$ which is defined by

$$(P_h\mathbf{u}, \mathbf{v}) = (\mathbf{u}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h.$$

We can then prove (cf. [14,10])

$$\|\mathbf{u} - P_h\mathbf{u}\| + h\|\mathbf{u} - P_h\mathbf{u}\|_{1,h} \le ch^2\|\mathbf{u}\|_2, \quad \forall \mathbf{u} \in \mathbf{H}^2 \cap \mathbf{H}_0^1. \tag{3.18}$$

3.3 Discrete inequalities

To handle the nonlinear convection term, we define

$$B(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{grad}\mathbf{v} + \frac{1}{2}(\mathrm{div}\mathbf{u})\mathbf{v}, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{H}_0^1,$$
$$b(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (B(\mathbf{u}, \mathbf{v}), \mathbf{w}) = \frac{1}{2}[(\mathbf{u} \cdot \mathbf{grad}\mathbf{v}, \mathbf{w}) - (\mathbf{u} \cdot \mathbf{grad}\mathbf{w}, \mathbf{v})], \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}_0^1,$$
$$b_h(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \frac{1}{2}[(\mathbf{u} \cdot \mathbf{grad}_h\mathbf{v}, \mathbf{w}) - (\mathbf{u} \cdot \mathbf{grad}_h\mathbf{w}, \mathbf{v})], \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}_h.$$

An elementary calculation yields

$$(B(\mathbf{u}, \mathbf{v}), \mathbf{w}) - b_h(\mathbf{u}, \mathbf{v}, \mathbf{w}) = X_3(\mathbf{u}, \mathbf{v}, \mathbf{w}), \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{H}_0^1, \mathbf{w} \in \mathbf{H}.$$

We need to obtain the discrete Sobolev inequality on the finite element space. The basic idea is to consider reference triangle $\hat{T}$ (or tetrahedron) with vertices at $\{(0, 0), (0, 1), (1, 0)\}$ (or $\{(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0)\}$). The constant of Sobolev inequality on this reference element $\hat{T}$ is fixed. For any element T, the affine transformation $\mathcal{F} : \hat{T} \to T$ takes form $\mathcal{F}(\hat{x}) = J\hat{x} + b$. Now consider a function v which restricts to the element T, then $\hat{v} = v \circ \mathcal{F}$ is a function defined on the reference element $\hat{T}$. We also denote constant $\sigma = \frac{r_o}{r_i}$, where r_o is the radius of the smallest ball which contains T, r_i is the radius of the largest ball which is contained by T. Quasi-uniform mesh implies that σ is uniformly bounded. By an elementary calculation, we have (cf. [10])

$$|\det(J)| = \frac{\mathrm{meas}T}{\mathrm{meas}\hat{T}} \approx h^n, \quad \|J\| \approx h, \quad \|J^{-1}\| \approx h^{-1}, \tag{3.19}$$

where $\approx$ means same order, $\|J\|$ is the matrix norm induced by Euclidean norm of R^n. Then by changing variable, chain rule and applying (3.19), the following estimation is obtained.

$$|\hat{v}|_{s,p,\hat{T}} \approx h^{s-\frac{n}{p}}|v|_{s,p,T}, \quad s \ge 0, 1 \le p < \infty, \tag{3.20}$$
$$|\hat{v}|_{s,p,\hat{T}} \approx h^{s}|v|_{s,p,T}, \quad s \ge 0, p = \infty, \tag{3.21}$$

where $|\cdot|_{s,p,T}$ denote Sobolev space $W^{s,p}$ in reference element T. Then we will prove following discrete Poincaré inequality and Sobolev inequality ($n = 2$ or 3):

Lemma 3.3

$$\begin{aligned}
&\|\mathbf{w}\| \le c\|\mathbf{grad}_h\mathbf{w}\|, && \forall \mathbf{w}\in\mathbf{H},\\
&\|\mathbf{w}\|_{\mathbf{L}^6} \le c\|\mathbf{grad}_h\mathbf{w}\|, && \forall \mathbf{w}\in\mathbf{H},\\
&\|\mathbf{w}\|_{\mathbf{L}^3} \le c\|\mathbf{w}\|^{1/2}\|\mathbf{grad}_h\mathbf{w}\|^{1/2}, && \forall \mathbf{w}\in\mathbf{H},\\
&\|\mathbf{w}\|_{\mathbf{L}^\infty} \le c\|\mathbf{grad}_h\mathbf{w}\|^{1/2}\|\Delta_h\mathbf{w}\|^{1/2}, && \forall \mathbf{w}\in\mathbf{V}_h.
\end{aligned}$$

Proof The proof is based on a scaling argument (See [8,14]). Since we may use similar techniques in later section, we prove the last inequality in the 3D case to show the idea in details. Other inequalities can be done similarly. Firstly we have Sobolev inequality on the reference element:

$$\begin{aligned}
&\|\hat f\|^2_{L^\infty,\hat T} \le c\|\hat f\|_{1,\hat T}\|\hat f\|_{2,\hat T}, \quad \forall \hat f\in H^2,\\
&\|\hat f\|^2_{L^\infty,\hat T} \le c\|\mathbf{grad}\hat f\|_{\hat T}\|\Delta\hat f\|_{\hat T}, \quad \forall \hat f\in H^2\cap L^2_0 \text{ and } \partial\hat f/\partial x_i\in L^2_0,\\
&\|\hat f\|^2_{L^\infty,\hat T} \le c\|\hat f\|^2_{\hat T}, \quad \forall \hat f\in P_1(\hat T).
\end{aligned}$$

We denote functions $\bar f$ and $\check f$ to be piecewise constant and piecewise linear approximations of function f defined by $\int_T(f-\bar f)dx=0$ and $\int_T D^\alpha(f-\check f)dx=0$ for all $0\le|\alpha|\le 1$, respectively. Then we have the following inequality:

$$\|\check f\|_T \le c(\|f\|_T + h\|\mathbf{grad} f\|_T).$$

The proof is as follows. $\check f = a_0+\sum_{i=1}^n a_ix_i$, where $a_i=\frac{1}{|T|}\int_T\frac{\partial f}{\partial x_i}dx$, $a_0=\frac{1}{|T|}\int_T(f-\sum_{i=1}^n a_ix_i)dx$, and

$$\begin{aligned}
|T|a_i^2 &= \int_T\left|\frac{\partial\check f}{\partial x_i}\right|^2dx \le c\int_T\left|\frac{\partial f}{\partial x_i}\right|^2dx,\\
\|\check f\|^2_T &= \int_T\left(a_0+\sum_{i=1}^n a_ix_i\right)^2dx\\
&= \int_T\left(\frac{1}{|T|}\int_T\left(f(y)-\sum_{i=1}^n a_iy_i\right)dy+\sum_{i=1}^n a_ix_i\right)^2dx\\
&\le \frac{1}{|T|^2}\int_T|T|\int_T\left(f(y)-\sum_{i=1}^n a_i(x_i-y_i)\right)^2dydx\\
&\le \frac{n+1}{|T|}\int_T\int_T\left[f(y)^2+\sum_{i=1}^n a_i^2(x_i-y_i)^2\right]dydx
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{n+1}{|T|} \int_T \int_T \left[f(y)^2 + h^2 \sum_{i=1}^n a_i^2 \right] dy dx \\
&= (n+1) \int_T f^2 dx + h^2 (n+1) \sum_{i=1}^n |T| a_i^2 \\
&\leq c \left(\| f \|_T^2 + h^2 \| \mathbf{grad} f \|_T^2 \right).
\end{aligned}
$$

To prove the discrete Sobolev inequality, we introduce an auxiliary function $\mathbf{v}$ which solves elliptic equation $\Delta \mathbf{v} = \Delta_h \mathbf{w}$, where $\mathbf{v} \in \mathbf{H}^2 \cap \mathbf{H}_0^1$. Then

$$
(\mathbf{grad}_h \mathbf{w}, \mathbf{grad}_h \psi) = (\mathbf{grad}_h \mathbf{v}, \mathbf{grad}_h \psi) - X_1(\mathbf{v}, \psi), \quad \forall \psi \in \mathbf{V}_h.
$$

By choosing $\psi = \mathbf{w} - \imath_h \mathbf{v}$ and using elliptic duality argument, this equation implies $\|\mathbf{v} - \mathbf{w}\| + h \|\mathbf{v} - \mathbf{w}\|_{1,h} \leq ch^2 \|\Delta \mathbf{v}\| = ch^2 \|\Delta_h \mathbf{w}\|$. From the definition of $\Delta_h \mathbf{w}$ we can easily obtain another inverse inequality $\|\Delta_h \mathbf{w}\| \leq ch^{-1} \|\mathbf{grad}_h \mathbf{w}\|$. Then using triangle inequality we have

$$
\|\mathbf{w}\|_{L^\infty} \leq \|\mathbf{w} - \check{\mathbf{v}}\|_{L^\infty} + \|\check{\mathbf{v}} - \mathbf{v}\|_{L^\infty} + \|\mathbf{v}\|_{L^\infty}.
$$

The last term $\|\mathbf{v}\|_{L^\infty}$ can be bounded by

$$
\begin{aligned}
\|\mathbf{v}\|_{L^\infty}^2 &\leq c \|\mathbf{grad} \mathbf{v}\| \|\Delta \mathbf{v}\| \leq c \left(\|\mathbf{grad}_h \mathbf{w}\| + \|\mathbf{grad}_h (\mathbf{w} - \mathbf{v})\| \right) \|\Delta_h \mathbf{w}\| \\
&\leq c \|\mathbf{grad}_h \mathbf{w}\| \|\Delta_h \mathbf{w}\|.
\end{aligned}
$$

The first term can be estimated as follow,

$$
\begin{aligned}
\|\mathbf{w} - \check{\mathbf{v}}\|_{L^\infty} &= \sup_{T \in \mathcal{T}_h} \|\mathbf{w} - \check{\mathbf{v}}\|_{L^\infty, T} \leq c \sup_{T \in \mathcal{T}_h} \|\widehat{\mathbf{w} - \check{\mathbf{v}}}\|_{L^\infty, \hat{T}} \leq c \sup_{T \in \mathcal{T}_h} \|\widehat{\mathbf{w} - \check{\mathbf{v}}}\|_{\hat{T}} \\
&\leq ch^{-3/2} \sup_{T \in \mathcal{T}_h} \|\mathbf{w} - \check{\mathbf{v}}\|_T \leq ch^{-3/2} (\|\mathbf{w} - \mathbf{v}\| + h \|\mathbf{w} - \mathbf{v}\|_{1,h}).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\|\mathbf{v} - \check{\mathbf{v}}\|_{L^\infty}^2 &= \sup_{T \in \mathcal{T}_h} \|\mathbf{v} - \check{\mathbf{v}}\|_{L^\infty, T}^2 = \sup_{T \in \mathcal{T}_h} \|\widehat{\mathbf{v} - \check{\mathbf{v}}}\|_{L^\infty, \hat{T}}^2 \\
&\leq c \sup_{T \in \mathcal{T}_h} |\widehat{\mathbf{v} - \check{\mathbf{v}}}|_{1,2,\hat{T}} |\widehat{\mathbf{v} - \check{\mathbf{v}}}|_{2,2,\hat{T}} \\
&\leq c \sup_{T \in \mathcal{T}_h} h^{-1/2} |\mathbf{v} - \check{\mathbf{v}}|_{1,2,T} h^{1/2} |\mathbf{v} - \check{\mathbf{v}}|_{2,2,T} \leq c \sup_{T \in \mathcal{T}_h} (|\mathbf{v}|_{1,2,T} + |\check{\mathbf{v}}|_{1,2,T}) |\mathbf{v}|_{2,2,T} \\
&\leq c \sup_{T \in \mathcal{T}_h} \|\mathbf{grad} \mathbf{v}\| |\mathbf{v}|_{2,2,T} \leq c \|\mathbf{grad} \mathbf{v}\| \|\Delta \mathbf{v}\|.
\end{aligned}
$$

Combining them together and using the inverse estimate, we obtain the inequality. □

4 Error estimate of the finite element approximation

We now consider the finite element spatial discretization. Let function $\mathbf{u}_h \in \mathbf{V}_h$ and satisfy

$$(\mathbf{u}_{h,t}, \mathbf{v}) + a_h(\mathbf{u}_h, \mathbf{v}) + b_h(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}) = (\mathbf{f}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{4.1}$$
$$\mathbf{u}_h(0) = P_h \mathbf{u}_0. \tag{4.2}$$

Here we choose the initial velocity approximation as an $\mathbf{L}^2$ projection $P_h\mathbf{u}_0$ for technical reasons. We believe that the conclusion should be independent of its choice (for instance, an interpolation approximation should work as well). The existence and uniqueness for the solutions of both continuous equation (2.3) and discrete equations (4.1) are stated in [25,26]. To estimate the error and to deal with the nonlinear term, we split the error function $\mathbf{u} - \mathbf{u}_h$ to two parts by introducing an auxiliary function $\mathbf{u}^* \in \mathbf{V}_h$ which satisfies

$$(\mathbf{u}^*_t, \mathbf{v}) + a_h(\mathbf{u}^*, \mathbf{v}) + b_h(\mathbf{u}, \mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{4.3}$$
$$\mathbf{u}^*(0) = P_h \mathbf{u}_0. \tag{4.4}$$

The only difference between Eqs. (4.1) and (4.3) is the trilinear term b_h. Denote $\mathbf{e}^1 = \mathbf{u} - \mathbf{u}^*$, $\mathbf{e}^2 = \mathbf{u}^* - \mathbf{u}_h$, then $\mathbf{e}^1$ and $\mathbf{e}^2$ satisfy equations

$$(\mathbf{e}^1_t, \mathbf{v}) + a_h(\mathbf{e}^1, \mathbf{v}) = X_1(\mathbf{u}, \mathbf{v}) + \frac{1}{\epsilon} X_2(\mathrm{div}\mathbf{u}, \mathbf{v}) - X_3(\mathbf{u}, \mathbf{u}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{4.5}$$
$$\mathbf{e}^1(0) = \mathbf{u}_0 - P_h \mathbf{u}_0 \tag{4.6}$$

and

$$(\mathbf{e}^2_t, \mathbf{v}) + a_h(\mathbf{e}^2, \mathbf{v}) + b_h(\mathbf{u}, \mathbf{u}, \mathbf{v}) - b_h(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}) = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{4.7}$$
$$\mathbf{e}^2(0) = \mathbf{0}, \tag{4.8}$$

respectively.

4.1 Finite element approximation to the velocity

Lemma 4.1 *We define an operator S_h from $\mathbf{H}^2 \cap \mathbf{H}^1_0$ to $\mathbf{V}_h$ by*

$$a_h(S_h\mathbf{u} - \mathbf{u}, \mathbf{v}) = -X_1(\mathbf{u}, \mathbf{v}) - \frac{1}{\epsilon} X_2(\mathrm{div}\mathbf{u}, \mathbf{v}) + X_3(\mathbf{u}, \mathbf{u}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h. \tag{4.9}$$

Then we have

$$\|\mathbf{u}-S_h\mathbf{u}\|_{a,h}^2 \le ch^2\left(\|\mathbf{u}\|_2^2+\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1^2+\|\mathbf{u}\|_2^2\|\mathbf{u}\|_1^2\right), \tag{4.10}$$

$$\|\mathbf{u}-S_h\mathbf{u}\|+h\|\mathbf{u}-S_h\mathbf{u}\|_{1,h} \le ch^2\left(\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1+\|\mathbf{u}\|_2+\|\mathbf{u}\|_1\|\mathbf{u}\|_2\right), \tag{4.11}$$

$$\|\mathbf{u}_t-S_h\mathbf{u}_t\| \le ch^2\left(\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}_t\|_1+\|\mathbf{u}_t\|_2+\|\mathbf{u}\|_1\|\mathbf{u}_t\|_2+\|\mathbf{u}_t\|_1\|\mathbf{u}\|_2\right). \tag{4.12}$$

Proof The technique of the proof is quite standard. Define $\mathbf{z}^1=\mathbf{u}-\imath_h\mathbf{u}$, $\mathbf{z}^2=\imath_h\mathbf{u}-S_h\mathbf{u}$, then the error function $\mathbf{e}=\mathbf{u}-S_h\mathbf{u}=\mathbf{z}^1+\mathbf{z}^2$. Choosing $\mathbf{v}=\mathbf{z}^2$ in Eq. (4.9), we have

$$\frac{1}{\epsilon}\|\mathrm{div}_h\mathbf{z}^2\|^2+\|\mathbf{grad}_h\mathbf{z}^2\|^2=-\frac{1}{\epsilon}(\mathrm{div}_h\mathbf{z}^1,\mathrm{div}_h\mathbf{z}^2)$$
$$-(\mathbf{grad}_h\mathbf{z}^1,\mathbf{grad}_h\mathbf{z}^2)-X_1(\mathbf{u},\mathbf{z}^2)-\frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{u},\mathbf{z}^2)+X_3(\mathbf{u},\mathbf{u},\mathbf{z}^2).$$

By using inequalities (3.6), (3.7), (3.8), we obtain

$$\frac{1}{\epsilon}\|\mathrm{div}_h\mathbf{z}^2\|^2+\|\mathbf{grad}_h\mathbf{z}^2\|^2\le c\Bigg[\frac{1}{\epsilon}\|\mathrm{div}_h\mathbf{z}^1\|^2+\|\mathbf{grad}_h\mathbf{z}^1\|^2$$
$$+h^2\left(\|\mathbf{u}\|_2^2+\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1^2+\|\mathbf{u}\|_1^2\|\mathbf{u}\|_2^2\right)\Bigg].$$

From the definition of $\mathbf{z}^1$ and inequalities (3.11), (3.12), we have (4.10). To obtain $\mathbf{L}^2$ estimation, we need to consider a duality problem

$$A_\epsilon\mathbf{w}=\mathbf{e}, \tag{4.13}$$
$$\mathbf{w}|_{\partial\Omega}=\mathbf{0}. \tag{4.14}$$

A priori estimation gives

$$\frac{1}{\epsilon}\|\mathrm{div}\mathbf{w}\|_1+\|\mathbf{w}\|_2\le c\|\mathbf{e}\|.$$

We multiply $\mathbf{e}$ to both sides of Eq. (4.13) and choose $\mathbf{v}=\imath_h\mathbf{w}$ in Eq. (4.9), add them together to have

$$\|\mathbf{e}\|^2=-\frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{w},\mathbf{e})-X_1(\mathbf{w},\mathbf{e})+\frac{1}{\epsilon}(\mathrm{div}_h\mathbf{e},\mathrm{div}_h(\mathbf{w}-\imath_h\mathbf{w}))$$
$$+(\mathbf{grad}_h\mathbf{e},\mathbf{grad}_h(\mathbf{w}-\imath_h\mathbf{w}))+X_1(\mathbf{u},\imath_h\mathbf{w})+\frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{u},\imath_h\mathbf{w})-X_3(\mathbf{u},\mathbf{u},\imath_h\mathbf{w}).$$

We will estimate every term by invoking inequality (3.6)–(3.12).

$$\begin{aligned}
&|X_2(\mathrm{div}\mathbf{w}, \mathbf{e})| && \le ch\|\mathrm{div}\mathbf{w}\|_1\|\mathbf{e}\|_{1,h}, \\
&|X_1(\mathbf{w}, \mathbf{e})| && \le ch\|\mathbf{w}\|_2\|\mathbf{e}\|_{1,h}, \\
&|X_1(\mathbf{u}, \iota_h\mathbf{w})| && \le ch\|\mathbf{u}\|_2\|\iota_h\mathbf{w} - \mathbf{w}\|_{1,h} && \le ch^2\|\mathbf{u}_2\|\|\mathbf{w}\|_2, \\
&|X_2(\mathrm{div}\mathbf{u}, \iota_h\mathbf{w})| && \le ch\|\mathrm{div}\mathbf{u}\|_1\|\iota_h\mathbf{w} - \mathbf{w}\|_{1,h} && \le ch^2\|\mathrm{div}\mathbf{u}\|_1\|\mathbf{w}\|_2, \\
&|X_3(\mathbf{u}, \mathbf{u}, \iota_h\mathbf{w})| && \le ch^2\|\mathbf{u}\|_1\|\mathbf{u}\|_2\|\mathbf{w}\|_2, \\
&(\mathrm{div}_h\mathbf{e}, \mathrm{div}_h(\mathbf{w} - \iota_h\mathbf{w})) && \le ch\|\mathrm{div}_h\mathbf{e}\|\|\mathrm{div}\mathbf{w}\|_1, \\
&(\mathbf{grad}_h\mathbf{e}, \mathbf{grad}_h(\mathbf{w} - \iota_h\mathbf{w})) && \le ch\|\mathbf{grad}_h\mathbf{e}\|\|\mathbf{w}\|_2.
\end{aligned}$$

Then we use Young's inequality to obtain

$$\begin{aligned}
\|\mathbf{e}\|^2 \le \delta &\left(\frac{1}{\epsilon^2}\|\mathrm{div}\mathbf{w}\|_1^2 + \|\mathbf{w}\|_2^2\right) \\
&+ c_\delta\left(h^2\|\mathbf{e}\|_{1,h}^2 + h^4\left(\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|^2 + \|\mathbf{u}\|_2^2 + \|\mathbf{u}\|_1^2\|\mathbf{u}\|_2^2\right)\right).
\end{aligned}$$

Choosing a properly small δ and noting (4.10) we conclude (4.11). Now by taking time derivative of Eq. (4.9), and denoting $\mathbf{z} = \mathbf{u}_t - (S_h\mathbf{u})_t$, we have

$$a_h(\mathbf{z}, \mathbf{v}) = X_1(\mathbf{u}_t, \mathbf{v}) + \frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{u}_t, \mathbf{v}) - X_3(\mathbf{u}_t, \mathbf{u}, \mathbf{v}) - X_3(\mathbf{u}, \mathbf{u}_t, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h.$$

By similar argument, let $\mathbf{v} = S_h\mathbf{u}_t - \iota_h\mathbf{u}_t$, we can obtain

$$\|\mathbf{z}\|_{a,h}^2 \le ch^2\left(\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}_t\|_1^2 + \|\mathbf{u}_t\|_2^2 + \|\mathbf{u}_t\|_2^2\|\mathbf{u}\|_1^2 + \|\mathbf{u}\|_2^2\|\mathbf{u}_t\|_1^2\right).$$

Then the $\mathbf{L}^2$ estimation (4.12) follows from the duality argument. □

From Lemma 4.1 and a priori estimation Lemma 2.2 we can easily conclude

$$\begin{aligned}
&\|\mathbf{u} - S_h\mathbf{u}\| + h\|\mathbf{u} - S_h\mathbf{u}\|_{1,h} \le ch^2(M_2 + M_2M_1), \\
&\left[\int_0^T t(\|\mathbf{u}_t - (S_h\mathbf{u})_t\|^2 + h^2\|\mathbf{u}_t - (S_h\mathbf{u})_t\|_{1,h}^2)dt\right]^{1/2} \le ch^2M_3(1 + M_2 + M_1).
\end{aligned}$$

To simplify the constant notation, we define constant M as

$$M = 1 + \sum_{i=1}^{3} M_i^2. \tag{4.15}$$

Lemma 4.2

$$\left[\int_0^t \|\mathbf{e}^1\|^2 d\tau\right]^{1/2} \le cMh^2\sqrt{t}.$$

Proof We define two functions $\mathbf{f}_1 \in \mathbf{L}^2$ and $\mathbf{f}_2 \in \mathbf{V}_h$ as

$$\mathbf{f}_1 = \mathbf{f} - B(\mathbf{u}, \mathbf{u}), \quad (\mathbf{f}_2, \mathbf{w}) = (\mathbf{f}, \mathbf{w}) - b_h(\mathbf{u}, \mathbf{u}, \mathbf{w}), \forall \mathbf{w} \in \mathbf{V}_h,$$

and two linear operators $T : \mathbf{L}^2 \to \mathbf{H}_0^1$ and $T_h : \mathbf{L}^2 \to \mathbf{V}_h$ by

$$\frac{1}{\epsilon}(\mathrm{div}(T\mathbf{f}), \mathrm{div}\mathbf{v}) + (\mathbf{grad}(T\mathbf{f}), \mathbf{grad}\mathbf{v}) = (\mathbf{f}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{H}_0^1, \tag{4.16}$$

$$\frac{1}{\epsilon}(\mathrm{div}_h(T_h\mathbf{f}), \mathrm{div}_h\mathbf{v}) + (\mathbf{grad}_h(T_h\mathbf{f}), \mathbf{grad}_h\mathbf{v}) = (\mathbf{f}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h. \tag{4.17}$$

Evidently, T and T_h are well defined. Moreover, from the definition of T, T_h and $\mathbf{f}_i$, following the argument in Lemma 4.1, we have (the first inequality can also be found in [4])

$$\|(T - T_h)\mathbf{f}\| \le ch^2\|\mathbf{f}\|, \tag{4.18}$$

$$\|T\mathbf{f}_1 - T_h\mathbf{f}_2\| \le ch^2\left(\|\mathbf{u}\|_2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1 + \|\mathbf{u}\|_2\|\mathbf{u}\|_1\right). \tag{4.19}$$

This result is shown in Lemma 4.3. It can be verified that

$$2(T_h\mathbf{f}_t, \mathbf{f}) = \frac{d}{dt}(T_h\mathbf{f}, \mathbf{f}), \quad (T_h\mathbf{f}, \mathbf{f}) \ge 0, \quad \forall\, \mathbf{f} \in \mathbf{V}_h. \tag{4.20}$$

By definition,

$$(T_h\mathbf{f}, \mathbf{f}) = \frac{1}{\epsilon}(\mathrm{div}_h(T_h\mathbf{f}), \mathrm{div}_h(T_h\mathbf{f})) + (\mathbf{grad}_h(T_h\mathbf{f}), (T_h\mathbf{f})) \ge 0.$$

Since $\frac{d}{dt}$ and spatial derivatives div_h, $\mathbf{grad}_h$ are commute, we have

$$\begin{aligned}\frac{d}{dt}(T_h\mathbf{f}, \mathbf{f}) &= \frac{d}{dt}\left[\frac{1}{\epsilon}(\mathrm{div}_h(T_h\mathbf{f}), \mathrm{div}_h(T_h\mathbf{f})) + (\mathbf{grad}_h(T_h\mathbf{f}), \mathbf{grad}_h(T_h\mathbf{f}))\right]\\ &= 2((T_h\mathbf{f})_t, \mathbf{f}).\end{aligned}$$

Finally taking time derivative for Eq. (4.17), we have $(T_h\mathbf{f})_t = T_h\mathbf{f}_t$ and hence the result in (4.20). We rewrite Eqs. (2.3) and (4.3) as

$$\begin{aligned}\mathbf{u} &= T(\mathbf{f}_1 - \mathbf{u}_t),\\ \mathbf{u}^* &= T_h(\mathbf{f}_2 - \mathbf{u}_t^*).\end{aligned}$$

Hence

$$\begin{aligned}T_h\mathbf{e}_t^1 + \mathbf{e}^1 &= T_h(\mathbf{u}_t - \mathbf{u}_t^*) + (\mathbf{u} - \mathbf{u}^*)\\ &= (T_h - T)\mathbf{u}_t + T\mathbf{u}_t + \mathbf{u} - T_h\mathbf{f}_2\\ &= (T_h - T)\mathbf{u}_t + (T\mathbf{f}_1 - T_h\mathbf{f}_2).\end{aligned}$$

We take inner product with $\mathbf{e}^1$ for above equation, hence

$$\frac{1}{2}\frac{d}{dt}(T_h\mathbf{e}^1,\mathbf{e}^1)+\|\mathbf{e}^1\|^2=((T_h-T)\mathbf{u}_t+(T\mathbf{f}_1-T_h\mathbf{f}_2),\mathbf{e}^1).$$

After the integration from 0 to t, it follows

$$\begin{aligned}&(T_h\mathbf{e}^1(t),\mathbf{e}^1(t))+\int_0^t\|\mathbf{e}^1\|^2d\tau\\&\quad\le(T_h\mathbf{e}^1(0),\mathbf{e}^1(0))+4\int_0^t(\|(T_h-T)\mathbf{u}_t\|^2+\|T\mathbf{f}_1-T_h\mathbf{f}_2\|^2)d\tau.\end{aligned}$$

We notice that $T_h\mathbf{e}^1(0)=T_h(\mathbf{u}_0-P_h\mathbf{u}_0)$, and by the definition of the operator P_h,

$$(P_h\mathbf{u}_0-\mathbf{u}_0,\mathbf{v})=0,\quad\forall\mathbf{v}\in\mathbf{V}_h,$$

which implies $(T_h\mathbf{e}^1(0),\mathbf{e}^1(0))=0$. Combining it with Lemma 4.3 below, and noting T_h is positive definite, we complete the proof. □

Lemma 4.3 *Using the notation in Lemma 4.2, we have*

$$\begin{aligned}&\|(T-T_h)\mathbf{f}\|\le ch^2\|\mathbf{f}\|,\\&\|T\mathbf{f}_1-T_h\mathbf{f}_2\|\le ch^2\left(\|\mathbf{u}\|_2+\frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1+\|\mathbf{u}\|_2\|\mathbf{u}\|_1\right).\end{aligned}$$

Proof We start from the first claim. Suppose $\mathbf{v}=T\mathbf{f}$, $\mathbf{v}_h=T_h\mathbf{f}$ and $\mathbf{e}=\mathbf{v}-\mathbf{v}_h$, then $\mathbf{e}$ satisfies equation

$$a_h(\mathbf{e},\mathbf{w})=X_1(\mathbf{v},\mathbf{w})+\frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{v},\mathbf{w}),\quad\forall\mathbf{w}\in\mathbf{V}_h.$$

Similar to Lemma 4.1, we let $\mathbf{w}=\mathbf{v}_h-\imath_h\mathbf{v}$, then the $\mathbf{L}^2$ estimate follows from the standard duality argument. The second result can be proved by a similar approach. Let $\mathbf{v}=T\mathbf{f}_1$, $\mathbf{v}_h=T_h\mathbf{f}_2$ and $\mathbf{e}=\mathbf{v}-\mathbf{v}_h$. Then $\mathbf{e}$ satisfies

$$a_h(\mathbf{e},\mathbf{w})=X_1(\mathbf{v},\mathbf{w})+\frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{v},\mathbf{w})-X_3(\mathbf{u},\mathbf{u},\mathbf{w}),\quad\forall\mathbf{w}\in\mathbf{V}_h.$$

Again, the result follows from the standard duality argument. □

Lemma 4.4

$$\begin{aligned}&\|\mathbf{e}^1\|+h\|\mathbf{e}^1\|_{1,h}\le ch^2M,\\&\int_0^t\tau\|S_h\mathbf{u}-\mathbf{u}^*\|^2_{a,h}d\tau\le cth^4M^2.\end{aligned}$$

Proof We have $\mathbf{e}^1 = \mathbf{u} - S_h\mathbf{u} + S_h\mathbf{u} - \mathbf{u}^*$, then rewrite the Eq. (4.5) as

$$(\mathbf{e}_t^1, \mathbf{v}) + a_h(S_h\mathbf{u} - \mathbf{u}^*, \mathbf{v}) = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h. \tag{4.21}$$

Choosing $\mathbf{v} = S_h\mathbf{u} - \mathbf{u}^*$ we have

$$\frac{1}{2}\frac{d}{dt}\|S_h\mathbf{u} - \mathbf{u}^*\|^2 + \|S_h\mathbf{u} - \mathbf{u}^*\|_{a,h}^2 \le (S_h\mathbf{u} - \mathbf{u}^*, \mathbf{u}_t - (S_h\mathbf{u})_t).$$

We multiply it by $2t$, integrate from 0 to t and obtain

$$\begin{aligned} t\|S_h\mathbf{u} - \mathbf{u}^*\|^2 &+ \int_0^t 2\tau\|S_h\mathbf{u} - \mathbf{u}^*\|_{a,h}^2 d\tau \\ &\le \int_0^t [\|S_h\mathbf{u} - \mathbf{u}^*\|^2 + 2\tau|(S_h\mathbf{u} - \mathbf{u}^*, \mathbf{u}_t - (S_h\mathbf{u})_t)|]d\tau \\ &\le \int_0^t [2\|S_h\mathbf{u} - \mathbf{u}^*\|^2 + \tau^2\|\mathbf{u}_t - (S_h\mathbf{u})_t)\|^2]d\tau \\ &\le 4\int_0^t (\|S_h\mathbf{u} - \mathbf{u}\|^2 + \|\mathbf{e}^1\|^2)d\tau + t\int_0^t \tau\|\mathbf{u}_t - (S_h\mathbf{u})_t)\|^2 d\tau. \end{aligned}$$

Then using Lemmas 4.1, 4.2 and 2.2, we have

$$t\|S_h\mathbf{u} - \mathbf{u}^*\|^2 \le cth^4M^2.$$

Combining it with inverse estimation and Lemma 4.1 yields

$$\|\mathbf{u} - \mathbf{u}^*\| + h\|\mathbf{u} - \mathbf{u}^*\|_{1,h} \le ch^2M.$$

□

To obtain the error estimation of $\mathbf{e}^2$, we should handle the nonlinear convection term. The following a priori estimation for $\mathbf{u}^*$ is needed.

Lemma 4.5

$$\|\mathbf{u}^*\|_{\mathbf{L}^\infty}^2 + \|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3} \le c\|\mathbf{grad}\mathbf{u}\|^{1/2}\|\Delta\mathbf{u}\|^{1/2} + Mh^{1/2}.$$

Proof The proof is similar to Lemma 3.3. By the triangle inequality and the notation in Sect. 3, we have

$$\begin{aligned} \|\mathbf{u}^*\|_{\mathbf{L}^\infty} &\le \|\mathbf{u}^* - \check{\mathbf{u}}\|_{\mathbf{L}^\infty} + \|\mathbf{u} - \check{\mathbf{u}}\|_{\mathbf{L}^\infty} + \|\mathbf{u}\|_{\mathbf{L}^\infty}, \\ \|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3} &\le \|\mathbf{grad}_h(\mathbf{u}^* - \check{\mathbf{u}})\|_{\mathbf{L}^3} + \|\mathbf{grad}_h(\check{\mathbf{u}} - \mathbf{u}^*)\|_{\mathbf{L}^3} + \|\mathbf{grad}_h\mathbf{u}\|_{\mathbf{L}^3}. \end{aligned}$$

Following the argument in Lemma 3.3, we have

$$\|\mathbf{u}^* - \check{\mathbf{u}}\|_{\mathbf{L}^\infty} \le ch^{-3/2}(\|\mathbf{u}^* - \mathbf{u}\| + h\|\mathbf{u}^* - \mathbf{u}\|_{1,h}),$$
$$\|\mathbf{u} - \check{\mathbf{u}}\|_{\mathbf{L}^\infty} \le c\|\mathbf{gradu}\|^{1/2}\|\Delta\mathbf{u}\|^{1/2},$$
$$\|\mathbf{grad}_h(\mathbf{u}^* - \check{\mathbf{u}})\|_{\mathbf{L}^3} \le ch^{-1/2}\|\mathbf{grad}_h(\mathbf{u} - \mathbf{u}^*)\|,$$
$$\|\mathbf{grad}_h(\check{\mathbf{u}} - \mathbf{u}^*)\|_{\mathbf{L}^3} \le c\|\mathbf{gradu}\|^{1/2}\|\Delta\mathbf{u}\|^{1/2}.$$

This combined with (continuous) Sobolev inequality and Lemma 4.4 yields the result of the lemma. □

Lemma 4.6 *For all* $\mathbf{u} \in \mathbf{H}^2 \cap \mathbf{H}_0^1$, $\mathbf{v}, \mathbf{w} \in \mathbf{H}$, *we have*

$$|b_h(\mathbf{u}, \mathbf{w}, \mathbf{v})| + |b_h(\mathbf{u}, \mathbf{v}, \mathbf{w})| \le c\|\mathbf{u}\|_2\|\mathbf{grad}_h\mathbf{v}\|(\|\mathbf{w}\| + h\|\mathbf{grad}_h\mathbf{w}\|).$$

Proof Integrating by parts at each element yields

$$b_h(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (\mathbf{u} \cdot \mathbf{grad}_h\mathbf{v}, \mathbf{w}) + \frac{1}{2}((\mathrm{div}\mathbf{u})\mathbf{v}, \mathbf{w})$$
$$-X_3(\mathbf{u}, \mathbf{v}, \mathbf{w}), \quad \forall \mathbf{u} \in \mathbf{H}^2 \cap \mathbf{H}_0^1, \mathbf{v}, \mathbf{w} \in \mathbf{H}.$$

Applying Hölder inequality, Sobolev inequality and discrete Sobolev inequality we have

$$|(\mathbf{u} \cdot \mathbf{grad}_h\mathbf{v}, \mathbf{w})| \le \|\mathbf{u}\|_{\mathbf{L}^\infty}\|\mathbf{grad}_h\mathbf{v}\|\|\mathbf{w}\| \le c\|\mathbf{u}\|_2\|\mathbf{grad}_h\mathbf{v}\|\|\mathbf{w}\|,$$
$$|((\mathrm{div}\mathbf{u})\mathbf{v}, \mathbf{w})| \le \|\mathrm{div}\mathbf{u}\|_{L^6}\|\mathbf{v}\|_{\mathbf{L}^3}\|\mathbf{w}\| \le c\|\mathrm{div}\mathbf{u}\|_1\|\mathbf{grad}_h\mathbf{v}\|\|\mathbf{w}\|.$$

The lemma follows from the above inequalities, skew-symmetry of b_h and inequality (3.8). □

Lemma 4.7

$$\|\mathbf{e}^2\| \le cMe^{M^2t}\sqrt{t}h^2.$$

Proof Noting $\mathbf{u} - \mathbf{u}_h = \mathbf{e}^1 + \mathbf{e}^2$, we have

$$b_h(\mathbf{u}, \mathbf{u}, \mathbf{v}) - b_h(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v})$$
$$= b_h(\mathbf{u}, \mathbf{e}^1, \mathbf{v}) + b_h(\mathbf{u}_h, \mathbf{e}^2, \mathbf{v}) + b_h(\mathbf{e}^1, \mathbf{u}^*, \mathbf{v}) + b_h(\mathbf{e}^2, \mathbf{u}^*, \mathbf{v}).$$

Let $\mathbf{v} = \mathbf{e}^2$ in Eq. (4.7). Then we obtain

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{e}^2\|^2 + \frac{1}{\epsilon}\|\mathrm{div}_h\mathbf{e}^2\|^2 + \|\mathbf{grad}_h\mathbf{e}^2\|^2 = -(I_1 + I_2 + I_3),$$

where

$$I_1 = b_h(\mathbf{u}, \mathbf{e}^1, \mathbf{e}^2), \quad I_2 = b_h(\mathbf{e}^1, \mathbf{u}^*, \mathbf{e}^2), \quad I_3 = b_h(\mathbf{e}^2, \mathbf{u}^*, \mathbf{e}^2).$$

Then we will estimate I_i by using Lemma 4.6, Hölder inequality and discrete Sobolev inequality.

$$
\begin{aligned}
|I_1| &\le c\|\mathbf{u}\|_2\|\mathbf{grad}_h\mathbf{e}^2\|(\|\mathbf{e}^1\| + h\|\mathbf{e}^1\|_{1,h}), \quad \text{(see Lemma 4.6)},\\
|I_2| &\le |(\mathbf{e}^1\cdot\mathbf{grad}_h\mathbf{u}^*, \mathbf{e}^2)| + |(\mathbf{e}^1\cdot\mathbf{grad}_h\mathbf{e}^2, \mathbf{u}^*)|\\
&\le c(\|\mathbf{e}^1\|\|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3}\|\mathbf{e}^2\|_{\mathbf{L}^6} + \|\mathbf{e}^1\|\|\mathbf{grad}_h\mathbf{e}^2\|\|\mathbf{u}^*\|_{\mathbf{L}^\infty})\\
&\le c\|\mathbf{e}^1\|\|\mathbf{grad}_h\mathbf{e}^2\|(\|\mathbf{u}^*\|_{\mathbf{L}^\infty} + \|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3}),\\
|I_3| &\le |(\mathbf{e}^2\cdot\mathbf{grad}_h\mathbf{u}^*, \mathbf{e}^2)| + |(\mathbf{e}^2\cdot\mathbf{grad}_h\mathbf{e}^2, \mathbf{u}^*)|\\
&\le c(\|\mathbf{e}^2\|\|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3}\|\mathbf{e}^2\|_{\mathbf{L}^6} + \|\mathbf{e}^2\|\|\mathbf{grad}_h\mathbf{e}^2\|\|\mathbf{u}^*\|_{\mathbf{L}^\infty})\\
&\le c\|\mathbf{e}^2\|\|\mathbf{grad}_h\mathbf{e}^2\|(\|\mathbf{u}^*\|_{\mathbf{L}^\infty} + \|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3}).
\end{aligned}
$$

We cancel $\|\mathbf{grad}_h\mathbf{e}^2\|^2$ by invoking Young's inequality,

$$\frac{d}{dt}\|\mathbf{e}^2\|^2 \le cM^2\|\mathbf{e}^2\|^2 + ch^4M^4.$$

Applying the Gronwall inequality and noting $\mathbf{e}^2(0) = 0$, we complete the proof. □

Theorem 4.8 *We assume ϵ is small enough, i.e. $\epsilon c_1 \le 1$, where c_1 is as in Lemma 2.1, and condition (2.5)–(2.6) are satisfied. Then there exists a constant C which only depends on Ω, N_1 and N_2 such that*

$$\|\mathbf{u}(t,\cdot) - \mathbf{u}_h(t,\cdot)\| \le h^2C.$$

Proof If we define a constant $C = cMe^{M^2T}$, this theorem is nothing but a combination of Lemma 4.4 and 4.7. □

4.2 Error estimate for the divergence of the solution

The divergence of $\mathbf{u}$ presents pressure due to the Eq. (1.5). It is important to find an alternative bound for $\mathrm{div}_h(\mathbf{u} - \mathbf{u}_h)$ instead of crudely replacing it by an upper bound $\mathbf{grad}_h(\mathbf{u} - \mathbf{u}_h)$. The first bound of $\mathrm{div}_h(\mathbf{u} - \mathbf{u}_h)$ is as following.

Lemma 4.9

$$\|\mathrm{div}_h(\mathbf{u} - S_h\mathbf{u})\|^2 \le c\epsilon h^2M^2, \tag{4.22}$$

$$\|\mathrm{div}_h(S_h\mathbf{u} - \mathbf{u}^*)\|^2 \le c\epsilon t^{-1}h^4M^2, \tag{4.23}$$

$$\frac{1}{\epsilon}\|\mathrm{div}_h(\mathbf{u}^* - \mathbf{u}_h)\|^2 + \int_0^t \|(\mathbf{u}^* - \mathbf{u}_h)_t\|^2 d\tau \le Cth^2. \tag{4.24}$$

Proof: Inequality (4.22) has already been proved in (4.10). Let $\mathbf{v} = (S_h\mathbf{u} - \mathbf{u}^*)_t$ in Eq. (4.21). Hence

$$\|(S_h\mathbf{u} - \mathbf{u}^*)_t\|^2 + \frac{1}{2}\frac{d}{dt}\|S_h\mathbf{u} - \mathbf{u}^*\|_{a,h}^2 = ((S_h\mathbf{u} - \mathbf{u})_t, (S_h\mathbf{u} - \mathbf{u}^*)_t).$$

Multiplying it by t^2, and integrating it from 0 to t, we have

$$\begin{aligned} &t^2\|S_h\mathbf{u} - \mathbf{u}^*\|_{a,h}^2 + \int_0^t \tau^2\|(S_h\mathbf{u} - \mathbf{u}^*)_t\|^2 d\tau \\ &\quad\le \int_0^t \tau^2\|\mathbf{u}_t - (S_h\mathbf{u})_t\|^2 d\tau + 2\int_0^t \tau\|S_h\mathbf{u} - \mathbf{u}^*\|_{a,h}^2 \le cth^4M^2. \end{aligned} \tag{4.25}$$

The last step follows from Lemma 4.1 and Lemma 4.4. Dividing t^2 we obtain (4.23). To show the last inequality, we let $\mathbf{v} = (\mathbf{u}^* - \mathbf{u}_h)_t$ in Eq. (4.7),

$$\|\mathbf{e}_t^2\|^2 + \frac{d}{dt}\|\mathbf{e}^2\|_{a,h}^2 \le I_1 + I_2 + I_3 + I_4 + I_5,$$

where

$$\begin{aligned} &I_1 = |b_h(\mathbf{u}, \mathbf{e}^1, \mathbf{e}_t^2)|, \quad I_2 = |b_h(\mathbf{u}^*, \mathbf{e}^2, \mathbf{e}_t^2)|, \quad I_3 = |b_h(\mathbf{e}^2, \mathbf{e}^2, \mathbf{e}_t^2)|, \\ &I_4 = |b_h(\mathbf{e}^1, \mathbf{u}^*, \mathbf{e}_t^2)|, \quad I_5 = |b_h(\mathbf{e}^2, \mathbf{u}^*, \mathbf{e}_t^2)|. \end{aligned}$$

Combining discrete Sobolev inequality, Hölder inequality, inverse estimation, we have

$$\begin{aligned} I_1 &\le c\|\mathbf{u}\|_{\mathbf{L}^\infty}(\|\mathbf{e}_t^2\|\|\mathbf{e}^1\|_{1,h} + \|\mathbf{e}_t^2\|_{1,h}\|\mathbf{e}^1\|) \le c\|\mathbf{u}\|_{\mathbf{L}^\infty}\|\mathbf{e}_t^2\|(\|\mathbf{e}^1\|_{1,h} + h^{-1}\|\mathbf{e}^1\|), \\ I_2 &\le c\|\mathbf{u}^*\|_{\mathbf{L}^\infty}(\|\mathbf{e}_t^2\|\|\mathbf{e}^2\|_{1,h} + \|\mathbf{e}_t^2\|_{1,h}\|\mathbf{e}^2\|) \le c\|\mathbf{u}^*\|_{\mathbf{L}^\infty}\|\mathbf{e}_t^2\|(\|\mathbf{e}^2\|_{1,h} + h^{-1}\|\mathbf{e}^2\|), \\ I_3 &\le c\|\mathbf{e}^2\|_{\mathbf{L}^6}(\|\mathbf{e}^2\|_{1,h}\|\mathbf{e}_t^2\|_{\mathbf{L}^3} + \|\mathbf{e}_t^2\|_{1,h}\|\mathbf{e}^2\|_{\mathbf{L}^3}) \le ch^{-5/2}\|\mathbf{e}^2\|^2\|\mathbf{e}_t^2\|, \\ I_4 &\le c\|\mathbf{e}^1\|(\|\mathbf{e}_t^2\|_{\mathbf{L}^6}\|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3} + \|\mathbf{e}_t^2\|_{1,h}\|\mathbf{u}^*\|_{\mathbf{L}^\infty}) \\ &\le ch^{-1}\|\mathbf{e}^1\|\|\mathbf{e}_t^2\|(\|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3} + \|\mathbf{u}^*\|_{\mathbf{L}^\infty}), \\ I_5 &\le c\|\mathbf{e}^2\|(\|\mathbf{e}_t^2\|_{\mathbf{L}^6}\|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3} + \|\mathbf{e}_t^2\|_{1,h}\|\mathbf{u}^*\|_{\mathbf{L}^\infty}) \\ &\le ch^{-1}\|\mathbf{e}^2\|\|\mathbf{e}_t^2\|(\|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3} + \|\mathbf{u}^*\|_{\mathbf{L}^\infty}). \end{aligned}$$

Then we apply Lemma 4.4–4.7 and obtain

$$\|\mathbf{e}_t^2\|^2 + \frac{d}{dt}\|\mathbf{e}^2\|_{a,h}^2 \le Ch^2.$$

Then integrating it from 0 to t yields inequality (4.24). □

From Lemma 4.9 and triangle inequality, we immediately have $\|\text{div}_h(\mathbf{u}-\mathbf{u}_h)\| = \sqrt{\epsilon}O(h+t^{-1}h^2)$. Since the pressure in the penalized Navier-Stokes equations is given by $-\frac{1}{\epsilon}\text{div}\mathbf{u}$, but the error bound for $\text{div}\mathbf{u}$ we just derived is order $\sqrt{\epsilon}$. It implies that the error bound for pressure is still inversely proportional to $\sqrt{\epsilon}$. Next we would like to recover the order ϵ accuracy. We notice that div_h operator is from $\mathbf{V}_h$ to L_h, and $(\mathbf{V}_h, L_h)$ satisfies discrete inf-sup condition (3.1). To obtain the ϵ independent result, we may need more regularity requirement (i.e. condition (2.7)) besides the assumption (2.5)–(2.6) in Theorem 4.8. Similar to (4.15), we define a new constant $\bar{M}$ as

$$\bar{M} = 1 + \sum_{i=1}^{5} M_i^2. \tag{4.26}$$

The following two lemmas are required in later proof.

Lemma 4.10

$$\|\mathbf{e}_t^1\| \le ct^{-1}h^2M.$$

Proof Taking time derivative of (4.9) and following the argument of Lemma 4.1 we can obtain

$$\begin{aligned}&\|(\mathbf{u}-S_h\mathbf{u})_{tt}\| \\ &\le ch^2\left(\frac{1}{\epsilon}\|\text{div}\mathbf{u}_{tt}\|_1 + \|\mathbf{u}_{tt}\| + \|\mathbf{u}_{tt}\|_2\|\mathbf{u}\|_1 + \|\mathbf{u}_t\|_1\|\mathbf{u}_t\|_2 + \|\mathbf{u}\|_2\|\mathbf{u}_{tt}\|_1\right).\end{aligned}$$

Combing it with Lemma 2.2 we have

$$\int_0^t \tau^3\|(\mathbf{u}-S_h\mathbf{u})_{tt}\|^2 d\tau \le ch^4\bar{M}^2. \tag{4.27}$$

Taking time derivative of Eq. (4.21) and letting $\mathbf{v} = (S_h\mathbf{u}-\mathbf{u}^*)_t$ we have

$$\frac{d}{dt}\|(S_h\mathbf{u}-\mathbf{u}^*)_t\|^2 \le ((S_h\mathbf{u}-\mathbf{u})_{tt}, (S_h\mathbf{u}-\mathbf{u}^*)_t).$$

Then multiplying t^3 in both sides and taking integration from 0 to t, we obtain

$$\begin{aligned}t^3\|(S_h\mathbf{u}-\mathbf{u}^*)_t\|^2 &\le 4\int_0^t [\tau^2\|(S_h\mathbf{u}-\mathbf{u}^*)_t\|^2 + \tau^4\|(S_h\mathbf{u}-\mathbf{u})_{tt}\|^2]d\tau \\ &\le cth^4\bar{M}^2.\end{aligned}$$

Where the last step follows from (4.25) and (4.27). Dividing t^3 from both sides and applying Lemma 4.1 we obtain the result. □

Lemma 4.11 *Similar to Theorem 4.8, if we define constant* $\bar{C} = c\bar{M}e^{\bar{M}^2 T}$, *then we have*

$$\|\mathbf{e}_t^2\| \le \bar{C}^2(h + t^{-1/2}h^2).$$

Proof Taking the time derivative of Eq. (4.7), choosing $\mathbf{v} = \mathbf{e}_t^2$ and applying similar technique to control the nonlinear term, we have

$$\frac{d}{dt}\|\mathbf{e}_t^2\|^2 + \|\mathbf{e}_t^2\|_{a,h}^2 \le I,$$

where $I = |b_h(\mathbf{u}, \mathbf{u}_t, \mathbf{e}_t^2) + b_h(\mathbf{u}_t, \mathbf{u}, \mathbf{e}_t^2) - b_h(\mathbf{u}_h, \mathbf{u}_{h,t}, \mathbf{e}_t^2) - b_h(\mathbf{u}_{h,t}, \mathbf{u}_h, \mathbf{e}_t^2)|$. We substitute $\mathbf{u}_h = \mathbf{u} - \mathbf{e}^1 - \mathbf{e}^2$ into the formula of I and obtain

$$\begin{aligned}
I \le\ & |b_h(\mathbf{u}, \mathbf{e}_t^1, \mathbf{e}_t^2)| + |b_h(\mathbf{e}^2, \mathbf{u}_t, \mathbf{e}_t^2)| + |b_h(\mathbf{e}^1, \mathbf{u}_t, \mathbf{e}_t^2)| + |b_h(\mathbf{e}^2, \mathbf{e}_t^1, \mathbf{e}_t^2)| \\
& + |b_h(\mathbf{e}^1, \mathbf{e}_t^1, \mathbf{e}_t^2)| + |b_h(\mathbf{e}_t^2, \mathbf{u}, \mathbf{e}_t^2)| + |b_h(\mathbf{e}_t^1, \mathbf{u}, \mathbf{e}_t^2)| + |b_h(\mathbf{u}_t, \mathbf{e}^2, \mathbf{e}_t^2)| \\
& + |b_h(\mathbf{u}_t, \mathbf{e}^1, \mathbf{e}_t^2)| + |b_h(\mathbf{e}_t^2, \mathbf{e}^2, \mathbf{e}_t^2)| + |b_h(\mathbf{e}_t^2, \mathbf{e}^1, \mathbf{e}_t^2)| \\
& + |b_h(\mathbf{e}_t^1, \mathbf{e}^2, \mathbf{e}_t^2)| + |b_h(\mathbf{e}_t^1, \mathbf{e}^1, \mathbf{e}_t^2)|.
\end{aligned}$$

Applying Hölder inequality, discrete inequality, inverse estimate and Lemma 4.6 we obtain the following estimate for I,

$$\begin{aligned}
I \le c\Big(& \|\mathbf{u}_t\|_2(\|\mathbf{e}^1\| + \|\mathbf{e}^2\|)\|\mathbf{e}_t^2\|_{1,h} + \|\mathbf{u}\|_2(\|\mathbf{e}_t^1\| + \|\mathbf{e}_t^2\|)\|\mathbf{e}_t^2\|_{1,h} \\
& + h^{-1/2}\|\mathbf{e}_t^2\|(\|\mathbf{e}^2\|_{1,h} + \|\mathbf{e}^1\|_{1,h})\|\mathbf{e}_t^2\|_{1,h} + h^{-1/2}\|\mathbf{e}_t^1\|(\|\mathbf{e}^2\|_{1,h} + \|\mathbf{e}^1\|_{1,h})\|\mathbf{e}_t^2\|_{1,h}\Big).
\end{aligned}$$

Then using Young's inequality and applying the estimations in Lemma 4.4, 4.7 and 4.10, we have

$$I \le c\left(h^4\bar{C}^2\|\mathbf{u}_t\|_2^2 + t^{-2}h^4\bar{M}^3 + M\|\mathbf{e}_t^2\|^2 + h\bar{C}^2\|\mathbf{e}_t^2\|^2 + t^{-2}h^5\bar{C}^2\bar{M}^2\right).$$

Multiplying the inequality by t^2, taking integration from 0 to t and noting inequality (4.24), we obtain

$$t^2\|\mathbf{e}_t^2\|^2 \le \bar{C}^4(t^2h^2 + th^4),$$

which implies conclusion. □

Then the error estimation of $\mathrm{div}_h(\mathbf{u} - \mathbf{u}_h)$ can be immediately obtained from following lemmas.

Lemma 4.12

$$\|\mathrm{div}_h(\mathbf{u} - S_h\mathbf{u})\| \le c\epsilon h\bar{M}, \tag{4.28}$$

$$\|\mathrm{div}_h(\mathbf{u}^* - S_h\mathbf{u})\| \le c\epsilon\bar{M}(h + t^{-1}h^2), \tag{4.29}$$

$$\|\mathrm{div}_h(\mathbf{u}^* - \mathbf{u}_h)\| \le c\epsilon\bar{C}^2(h + t^{-1/2}h^2). \tag{4.30}$$

Proof We will prove these claims one by one. We denote $\mathbf{z} = \iota_h\mathbf{u} - S_h\mathbf{u}$. Due to the inf-sup condition (3.1) we can pick up a $\mathbf{v} \in \mathbf{V}_h$ such that

$$\|\mathrm{div}_h\mathbf{z}\| \le c\frac{|(\mathrm{div}_h\mathbf{z}, \mathrm{div}_h\mathbf{v})|}{\|\mathbf{grad}_h\mathbf{v}\|}.$$

Then from Eq. (4.9), we have

$$\begin{aligned}\frac{1}{\epsilon}\|\mathrm{div}_h\mathbf{z}\| &\le \frac{c}{\|\mathbf{grad}_h\mathbf{v}\|}\Big|\frac{1}{\epsilon}(\mathrm{div}_h(\iota_h\mathbf{u}-\mathbf{u}), \mathrm{div}_h\mathbf{v}) + (\mathbf{grad}_h(S_h\mathbf{u}-\mathbf{u}), \mathbf{grad}_h\mathbf{v}) \\ &\quad + X_1(\mathbf{u},\mathbf{v}) + \frac{1}{\epsilon}X_2(\mathrm{div}\mathbf{u},\mathbf{v}) - X_3(\mathbf{u},\mathbf{u},\mathbf{v})\Big| \\ &\le c\left[\frac{1}{\epsilon}\|\mathrm{div}_h(\iota_h\mathbf{u}-\mathbf{u})\| + \|\mathbf{u}-S_h\mathbf{u}\|_{1,h} + h\left(\|\mathbf{u}\|_2 + \frac{1}{\epsilon}\|\mathrm{div}\mathbf{u}\|_1 + \|\mathbf{u}\|_2\|\mathbf{u}\|_1\right)\right] \\ &\le ch\bar{M}.\end{aligned}$$

It implies (4.28). To prove (4.29), we need to choose a proper $\mathbf{v}$ in Eq. (4.21) such that

$$\begin{aligned}\tfrac{1}{\epsilon}\|\mathrm{div}_h(S_h\mathbf{u}-\mathbf{u}^*)\| &\le c\frac{1}{\epsilon}\frac{|(\mathrm{div}_h(S_h\mathbf{u}-\mathbf{u}^*), \mathrm{div}_h\mathbf{v})|}{\|\mathbf{grad}_h\mathbf{v}\|} \\ &= \frac{c}{\|\mathbf{grad}_h\mathbf{v}\|}|-(\mathbf{e}_t^1,\mathbf{v}) - (\mathbf{grad}_h(S_h\mathbf{u}-\mathbf{u}^*), \mathbf{grad}_h\mathbf{v})| \\ &\le c(\|\mathbf{e}_t^1\| + \|S_h\mathbf{u}-\mathbf{u}^*\|_{1,h})\end{aligned}$$

Thus (4.29) follows from Lemma 4.10. To show the last inequality (4.30), we pick up a proper $\mathbf{v}$ in Eq. (4.7) such that

$$\begin{aligned}\tfrac{1}{\epsilon}\|\mathrm{div}_h\mathbf{e}^2\| &\le c\frac{1}{\epsilon}\frac{|(\mathrm{div}_h\mathbf{e}^2, \mathrm{div}_h\mathbf{v})|}{\|\mathbf{grad}_h\mathbf{v}\|} \\ &= \frac{c}{\|\mathbf{grad}_h\mathbf{v}\|}|b_h(\mathbf{u},\mathbf{e}^1,\mathbf{v}) + b_h(\mathbf{u}^*,\mathbf{e}^2,\mathbf{v}) + b_h(\mathbf{e}^2,\mathbf{e}^2,\mathbf{v}) + b_h(\mathbf{e}^1,\mathbf{u}^*,\mathbf{v}) \\ &\quad + b_h(\mathbf{e}^2,\mathbf{u}^*,\mathbf{v}) - (\mathbf{e}_t^2,\mathbf{v}) - (\mathbf{grad}_h\mathbf{e}^2, \mathbf{grad}_h\mathbf{v})|.\end{aligned}$$

Then applying discrete Sobolev inequality, Hölder's inequality and Poincaré inequality we obtain

$$\begin{aligned}|b_h(\mathbf{u},\mathbf{e}^1,\mathbf{v})| &\le c\|\mathbf{u}\|_{\mathbf{L}^\infty}\|\mathbf{v}\|_{1,h}\|\mathbf{e}^1\|_{1,h}, \\ |b_h(\mathbf{u}^*,\mathbf{e}^2,\mathbf{v})| &\le c\|\mathbf{u}^*\|_{\mathbf{L}^\infty}\|\mathbf{v}\|_{1,h}\|\mathbf{e}^2\|_{1,h}, \\ |b_h(\mathbf{e}^2,\mathbf{e}^2,\mathbf{v})| &\le c\|\mathbf{e}^2\|_{1,h}^2\|\mathbf{v}\|_{1,h}, \\ |b_h(\mathbf{e}^1,\mathbf{u}^*,\mathbf{v})| &\le c\|\mathbf{e}^1\|\|\mathbf{v}\|_{1,h}(\|\mathbf{u}^*\|_{\mathbf{L}^\infty} + \|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3}), \\ |b_h(\mathbf{e}^2,\mathbf{u}^*,\mathbf{v})| &\le c\|\mathbf{e}^2\|\|\mathbf{v}\|_{1,h}(\|\mathbf{u}^*\|_{\mathbf{L}^\infty} + \|\mathbf{grad}_h\mathbf{u}^*\|_{\mathbf{L}^3}).\end{aligned}$$

Hence,

$$\frac{1}{\epsilon}\|\mathrm{div}_h \mathbf{e}^2\| \le c(\|\mathbf{u}\|_2\|\mathbf{e}^1\|_{1,h} + \|\mathbf{u}\|_2\|\mathbf{e}^2\|_{1,h} + \|\mathbf{e}^2\|_{1,h}^2 + \|\mathbf{e}_t^2\| + \|\mathbf{e}^2\|_{1,h}).$$

Then applying Lemma 4.4, 4.7, 4.11 and an inverse estimate we can conclude (4.30). □

Combining all the results in Lemma 4.12 and noting that the pressure p can be recovered by div$\mathbf{u}$, we have the following ϵ-uniform error estimate for the pressure:

Theorem 4.13 *Besides all assumption in Theorem 4.8, we also assume condition (2.7) is satisfied. Then we have*

$$\mathrm{div}_h(\mathbf{u}(t,\cdot) - \mathbf{u}_h(t,\cdot)) \le C\epsilon(h + t^{-1}h^2).$$

If we solve pressure as $p = -\frac{1}{\epsilon}\mathrm{div}\mathbf{u}$ *and* $p_h = -\frac{1}{\epsilon}\mathrm{div}_h\mathbf{u}_h$, *it is*

$$\|p_h(t,\cdot) - p(t,\cdot)\| \le C(h + t^{-1}h^2).$$

Remark 4.1 This paper focuses on the spatial discretization of the penalized Navier-Stokes equation in a convex polygon or polyhedron and a fixed time interval $[0, T]$. To our knowledge, the best result on the error estimate between the solution of original Navier-Stokes equation and the penalized Navier-Stokes equation is given in [23]. If we let $(\mathbf{u}_{NS}, p_{NS})$ be the solution of the original Navier-Stokes equation and $\mathbf{u}$ be the penalized solution, then the following is obtained in [23]:

$$\sqrt{t}\|\mathbf{u} - \mathbf{u}_{NS}\| + t\|\mathbf{u} - \mathbf{u}_{NS}\|_1 + \left(\int_0^T s^2\|p - p_{NS}\|^2 ds\right)^{1/2} \le C\epsilon, \tag{4.31}$$

where $p = -\frac{1}{\epsilon}\mathrm{div}\mathbf{u}$.

Combing Theorem 4.8, Theorem 4.13 and (4.31), we have the following error estimate between the solution of the original Navier-Stokes equation and the finite element approximation (4.1)–(4.2):

Theorem 4.14 *Suppose* ϵ *be small enough, i.e.* $\epsilon c_1 \le 1$, *where* c_1 *is as in Lemma 2.1 and assume (2.5)–(2.7) are satisfied. Let* $(\mathbf{u}_{NS}, p_{NS})$ *be solution to the Navier-Stokes equation (1.1)–(1.3), and* $\mathbf{u}_h$ *be the solution of (4.1)–(4.2), then we have*

$$\sqrt{t}\|\mathbf{u}_{NS}(t,\cdot) - \mathbf{u}_h(t,\cdot)\| + \left(\int_0^T s^2\|p_h(s,\cdot) - p_{NS}(s,\cdot)\|^2 ds\right)^{1/2} \le C(h+\epsilon).$$

References

1. Andreas, P.: Projection and Quasi-Compressibility Methods for Solving the Incompressible Navier-Stokes Equations. B. G. Teubner, Stuttgart (1997)
2. Bramble, J.H., Hilbert, S.R.: Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation. SIAM J. Numer. Anal. **7**, 113–124 (1970)
3. Brefort, B., Ghidaglia, J.M., Temam, R.: Attractors for the penalized Navier-Stokes equations. SIAM J. Math. Anal. **19**, 1–21 (1988)
4. Brenner, S.C., Sung, L.-Y.: Linear finite element methods for planar linear elasticity. Math. Comput. **59**, 321–338 (1992)
5. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Springer, New York (1994)
6. Brezzi, F., Pitkaranta, J.: On the stabilization of finite element approximation of the Stokes problems. In: Efficient Solution of Elliptic system, Notes on Numerical Fluids Mechanics, vol. 10, pp. 11–19 (1984)
7. Chorin, A.J.: Numerical solution of the Navier-Stokes equations. Math. Comput. **23**, 745–762 (1969)
8. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
9. Crouzeix, M., Raviart, P.A.: Comforming and nonconforming finite element methods for solving the stationary Stokes equations. RAIRO **R-3**, 33–75 (1973)
10. Ern, A., Guermond, J.L.: Theory and Practice of finite elements. Applied Mathematical Sciences Series. Springer, New York (2004)
11. Girault, V., Raviart, P.A.: Finite Element Method for Navier-Stokes Equations: Theory and Algorithm. Springer, Berlin (1987)
12. He, Y.: Optimal error estimate of the penalty finite element method for the time-dependent Navier-Stokes equations. Math. Comput. **74**, 1201–1216 (2005)
13. He, Y., Li, J., Yang, X.: Two-level penalized finite element methods for the stationary Navier-Stoke equations. Int. J. Inf. Syst. Sci. **2**, 131–143 (2006)
14. Heywood, J.G., Rannacher, R.: Finite element approximation of the nonstationary Navier-Stokes problem, I. Regularity of solutions and second-order error estimates for spatial discretization. SIAM J. Numer. Anal. **19**, 275–311 (1982)
15. Hughes, T.J.R., Liu, W.T., Brooks, A.J.: Finite element analysis of incompressible viscous flows by the penalty function formulation. J. Comput. Phys. **30**, 1–60 (1979)
16. Lin, P.: A sequential regularization methods for time-dependent incompressibel Navier-Stokes equations. SIAM J. Numer. Anal. **34**, 1051–1071 (1997)
17. Lin, P., Liu, J.G., Lu, X.L.: Long time numerical solution of the Navier-Stokes equations based on a sequential regularization formulation. SIAM J. Sci. Comput. **31**, 398–419 (2008)
18. Lin, P., Liu, C.: Simulation of singularity dynamics in liquid crystal flows: a C^0 finite element approach. J. Comput. Phys. **215**, 348–362 (2006)
19. Lin, P., Liu, C., Zhang, H.: An energy law preserving C^0 finite element scheme for simulating the kinematic effects in liquid crystal flow dynamics. J. Comput. Phys. **227**, 1411–1427 (2007)
20. Lu, X.: Error Analysis for incompressible viscous flow based on a sequential regularization formulation. Ph.D. Dissertation (2006)
21. Lu, X.L., Lin, P., Liu, J.G.: Analysis of a sequential regularization method for the unsteady Navier-Stokes equations. Math. Comput. **77**, 1467–1494 (2008)
22. Payne, L.E., Weinberger, H.F.: An optimal Poincaré inequality for convex domains. Arch. Ration. Mech. Anal. **5**, 286–292 (1960)
23. Shen, J.: On error estimates of the penalty method for the unsteady Navier-Stokes equations. SIAM J. Numer. Anal. **32**, 386–403 (1995)
24. Temam, R.: Navier-Stokes Equations. North-Holland, Amsterdam (1977)
25. Temam, R.: Une méthode d'approximation de la solution des équations de Navier-Stokes. Bull. Soc. Math. France **98**, 115–152 (1968)
26. Temam, R.: Sure L'approximation de la solution des equations de Navier-Stokes par la méthods des pas fractionnaires II. Arch. Ration. Mech. Anal. **33**, 377–385 (1969)