Ann. Appl. Math.
doi: 10.4208/aam.OA-2023-0016

Vol. **40**, No. 1, pp. 43-70
February 2024

# Error Analysis of the Nonconforming $P_1$ Finite Element Method to the Sequential Regularization Formulation for Unsteady Navier-Stokes Equations

Yanming Lai[1], Kewei Liang[2], Ping Lin[3], Xiliang Lu[4,*] and Qimeng Quan[1]

[1] *School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China*
[2] *School of Mathematical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China*
[3] *Division of Mathematics, University of Dundee, Dundee DD1 4HN, UK*
[4] *School of Mathematics and Statistics and Hubei Key Laboratory of Computational Science, Wuhan University, Wuhan 430072, China*

Received 9 May 2023; Accepted (in revised version) 6 August 2023

---

**Abstract.** In this paper we investigate the nonconforming $P_1$ finite element approximation to the sequential regularization method for unsteady Navier-Stokes equations. We provide error estimates for a full discretization scheme. Typically, conforming $P_1$ finite element methods lead to error bounds that depend inversely on the penalty parameter $\epsilon$. We obtain an $\epsilon$-uniform error bound by utilizing the nonconforming $P_1$ finite element method in this paper. Numerical examples are given to verify theoretical results.

**AMS subject classifications**: 65M60, 76D05

**Key words**: Navier-Stokes equations, error estimates, finite element method, stabilization method.

---

*Corresponding author.
*Emails:* `laiyanming@whu.edu.cn` (Y. Lai), `matlkw@zju.edu.cn` (K. Liang), `p.lin@dundee.ac.uk` (P. Lin), `xllv.math@whu.edu.cn` (X. Lu), `quanqm@whu.edu.cn` (Q. Quan)

# 1 Introduction

Let $\Omega$ be a bounded convex polygon domain of $\mathbb{R}^2$ or $\mathbb{R}^3$ and $\Gamma$ its boundary. We consider the unsteady Navier-Stokes equations for a viscous incompressible fluid in $\Omega\times[0,T]$:

$$\boldsymbol{u}_t-\nu\Delta\boldsymbol{u}+\boldsymbol{u}\cdot\nabla\boldsymbol{u}+\nabla p=\boldsymbol{f}_{ext}, \tag{1.1a}$$

$$\mathrm{div}\boldsymbol{u}=0, \tag{1.1b}$$

$$\boldsymbol{u}|_\Gamma=\boldsymbol{0}, \quad \boldsymbol{u}(\boldsymbol{x},0)=\boldsymbol{u}_0(\boldsymbol{x}). \tag{1.1c}$$

Here $\boldsymbol{u}(\boldsymbol{x},t)$ is the velocity of the fluid, $p$ the pressure acting on the fluid, $\boldsymbol{f}_{ext}$ the external force, $\boldsymbol{u}_0$ the initial velocity and $\nu$ the dynamic viscosity. The Eqs. (1.1a)-(1.1c) can be written as the equivalent system below:

$$\begin{aligned}&\boldsymbol{u}_t-\nu\Delta\boldsymbol{u}+B(\boldsymbol{u},\boldsymbol{u})+\nabla p=\boldsymbol{f}_{ext},\\&\mathrm{div}\boldsymbol{u}=0,\\&\boldsymbol{u}|_\Gamma=\boldsymbol{0}, \quad \boldsymbol{u}(\boldsymbol{x},0)=\boldsymbol{u}_0(\boldsymbol{x}),\end{aligned}$$

where

$$B(\boldsymbol{u},\boldsymbol{u})=\boldsymbol{u}\cdot\nabla\boldsymbol{u}+\frac{1}{2}(\mathrm{div}\boldsymbol{u})\boldsymbol{u}.$$

Eqs. (1.1a)-(1.1c) present a long-recognized difficulty for numerical solution due to the coupling of $\boldsymbol{u}$ and $p$ by the incompressible equation, where the pressure $p$ does not explicitly appear. This results in an index-2 differential algebraic system (cf. [5, 11]) and may cause temporal instability in maintaining the algebraic constraint (or the incompressible equation in the Navier-Stokes context). Hence, direct discretization is not recommended. To overcome this difficulty, several methods have been proposed, such as the projection method (cf. [8, 15]), penalty method (cf. [4, 14]), iterative penalty method for steady problems (cf. [7]), Baumgarte stabilization (cf. [3]), and sequential regularization method (SRM) [11]. The SRM is based on methods for solving differential algebraic equations (cf. [1, 2]) and can be understood as a combination of the penalty method and Baumgarte stabilization (see [13]). It reads as follows: given $p_0(\boldsymbol{x},t)$ the initial guess, for $s=1,2,\ldots$, solve

$$(\boldsymbol{u}_s)_t-\nu\Delta\boldsymbol{u}_s+B(\boldsymbol{u}_s,\boldsymbol{u}_s)+\nabla p_s=\boldsymbol{f}_{ext}, \tag{1.2a}$$

$$\mathrm{div}(\alpha_1(\boldsymbol{u}_s)_t+\alpha_2\boldsymbol{u}_s)=\epsilon(p_{s-1}-p_s), \tag{1.2b}$$

$$\boldsymbol{u}_s|_\Gamma=\boldsymbol{0}, \quad \boldsymbol{u}_s(\boldsymbol{x},0)=\boldsymbol{u}_0(\boldsymbol{x}), \tag{1.2c}$$

where $\alpha_1$ and $\alpha_2$ are nonnegative constants and $\epsilon$ a small penalty parameter. It has been showed that $u-u_s$ and $p-p_s=\mathcal{O}(\epsilon^s)$. In other words, unlike the penalty method $\epsilon$ is not necessarily chosen to be very small so as to have a better-posed system to solve at each iterate $s$. It is easy to see that the SRM formulation scheme decouples $\boldsymbol{u}$ and $p$. At the $s$-th step, we need to solve an equation of the form

$$(\boldsymbol{u}_s)_t - \frac{1}{\epsilon}\nabla(\alpha_1 \mathrm{div}((\boldsymbol{u}_s)_t + \alpha_2 \mathrm{div}\boldsymbol{u}_s) - \nu\Delta\boldsymbol{u}_s + B(\boldsymbol{u}_s,\boldsymbol{u}_s) = \boldsymbol{f}_s, \tag{1.3a}$$

$$\boldsymbol{u}_s|_\Gamma = \boldsymbol{0}, \quad \boldsymbol{u}_s(\boldsymbol{x},0) = \boldsymbol{u}_0, \tag{1.3b}$$

where $\boldsymbol{f}_s = \boldsymbol{f}_{ext} - \nabla p_{s-1}$. The existence theorem for solution of above equation is shown in [13]. Moreover, a time discretization scheme of $\{\boldsymbol{u}_s^n\}_{n=0}^N$

$$\frac{\boldsymbol{u}_s^n - \boldsymbol{u}_s^{n-1}}{\Delta t} - \frac{1}{\epsilon}\nabla\Big(\alpha_1 \mathrm{div}\frac{\boldsymbol{u}_s^n - \boldsymbol{u}_s^{n-1}}{\Delta t} + \alpha_2 \mathrm{div}\boldsymbol{u}_s^n\Big) - \nu\Delta\boldsymbol{u}_s^n + B(\boldsymbol{u}_s^{n-1},\boldsymbol{u}_s^n) = \boldsymbol{f}_s^n, \tag{1.4a}$$

$$\boldsymbol{u}_s^n|_\Gamma = \boldsymbol{0},\ n=1,2,\cdots,N; \quad \boldsymbol{u}_s^0 = \boldsymbol{u}_0, \tag{1.4b}$$

is studied in [13]. Here $N=\frac{T}{\Delta t}$, $\boldsymbol{f}_s^n(\boldsymbol{x}):=\boldsymbol{f}_s(\boldsymbol{x},n\Delta t)$, $n=0,\cdots,N$, and $B(\cdot,\cdot)$ is a bilinear form which is defined later. In [17], an error estimate for a space discretization scheme is also given. The scheme reads

$$\begin{aligned} &\big((\boldsymbol{u}_{sh})_t,\boldsymbol{v}_h\big) + \frac{1}{\epsilon}\big(\alpha_1 \mathrm{div}_h(\boldsymbol{u}_{sh})_t + \alpha_2 \mathrm{div}_h\boldsymbol{u}_{sh}, \mathrm{div}_h\boldsymbol{v}_h\big) \\ &\quad + \big(\nu\nabla_h\boldsymbol{u}_{sh},\nabla_h\boldsymbol{v}_h\big) + b\big(\boldsymbol{u}_{sh},\boldsymbol{u}_{sh},\boldsymbol{v}_h\big) = \big(\boldsymbol{f}_s,\boldsymbol{v}_h\big), \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \\ &\boldsymbol{u}_h|_\Gamma = \boldsymbol{0}, \quad \boldsymbol{u}_h(\boldsymbol{x},0) = P\boldsymbol{u}_0, \end{aligned}$$

where $\boldsymbol{V}_h$ is $P_1$ conforming finite element spaces, $P$ is a projection from $\boldsymbol{H}_0^1(\Omega)$ to $\boldsymbol{V}_h$ and $b(\cdot,\cdot,\cdot)$ is a trilinear form defined later. Unfortunately, the error bound of the SRM method is inversely proportional to the chosen penalty parameter $\epsilon$. This can be problematic if one wants to decrease $\epsilon$ to reduce the number of SRM iterations $s$ while still achieving a high degree of accuracy in approximating equations (1.1a)-(1.1c). In this paper, we propose using the nonconforming $P_1$ finite element space for spatial discretization, which provides an error bound that does not depend on $\epsilon$. Specifically, we study a full discretization scheme, given by Eq. (1.5),

$$\begin{aligned} &\Big(\frac{\boldsymbol{u}_{sh}^n - \boldsymbol{u}_{sh}^{n-1}}{\Delta t},\boldsymbol{v}_h\Big) + \frac{1}{\epsilon}(\alpha_1 \mathrm{div}_h\frac{\boldsymbol{u}_{sh}^n - \boldsymbol{u}_{sh}^{n-1}}{\Delta t} + \alpha_2 \mathrm{div}_h\boldsymbol{u}_{sh}^n, \mathrm{div}_h\boldsymbol{v}_h) + (\nu\nabla_h\boldsymbol{u}_{sh}^n,\nabla_h\boldsymbol{v}_h) \\ &\quad + b_h(\boldsymbol{u}_{sh}^{n-1},\boldsymbol{u}_{sh}^n,\boldsymbol{v}_h) = (\boldsymbol{f}_s^n,\boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \quad n=1,2,\cdots,N, \end{aligned} \tag{1.5a}$$

$$\boldsymbol{u}_h^0 = \iota_h\boldsymbol{u}_0, \tag{1.5b}$$

where $\boldsymbol{V}_h$ refers to the nonconforming $P_1$ finite element space and $\iota_h$ is the associated interpolation operator. We note that error estimates for the nonconforming $P_1$ element approximation have been previously provided in [6] and [12] for the planar linear elasticity and penalized Navier-Stokes equations, respectively.

The remainder of the paper is structured as follows. In Section 2, we introduce the notations and briefly review the nonconforming $P_1$ finite element method. In Section 3, we discuss issues related to our assumptions and provide error estimates for the time discretization scheme. In Section 4, we present error estimates for the full discretization scheme. In the final section, we present some numerical experiments and discuss their implications. Without loss of generality, we assume that $\alpha_1=1$, $\alpha_2=\alpha$, and $\nu=1$. It is important to note that the constant $C$ appearing in our estimates may vary depending on $\alpha$, $\Omega$, and $T$, but it does not depend on $\epsilon$ or the discretization parameters.

# 2 Preliminaries

## 2.1 Notations

Let $D$ be the standard Sobolev weak derivative operator and $C_0^\infty(\Omega)$ be the space of infinitely differentiable functions with compact support in $\Omega$. For any integer $m\geq 0$, we denote by $H^m(\Omega)$ and $H_0^1(\Omega)$ the standard Sobolev spaces of order $m$, with the norm given by

$$\|v\|_{H^m(\Omega)}=\left(\sum_{|\alpha|\leq m}\|D^\alpha v\|^2\right)^{\frac{1}{2}},\quad \forall v\in H^m(\Omega).$$

We usually rewrite the space $L^2(\Omega):=H^0(\Omega)$ if $m=0$. Denote the $L^2(\Omega)$ inner product by $(\cdot,\cdot)$, and the corresponding $L^2(\Omega)$ norm is given by $\|\cdot\|=\sqrt{(\cdot,\cdot)}$. For the multi-dimensional product space in $\mathbb{R}^d$ $(d=2,3)$, we define $\boldsymbol{H}^m(\Omega)=(H^m(\Omega))^d$, where the superscript $d$ indicates the standard Cartesian product. The spaces $\boldsymbol{L}^2(\Omega)$ and $\boldsymbol{H}_0^1(\Omega)$ are defined similarly.

To handle the nonlinear term, we define

$$\begin{aligned}
&B(\boldsymbol{u},\boldsymbol{v})=\boldsymbol{u}\cdot\nabla\boldsymbol{v}+\frac{1}{2}(\mathrm{div}\boldsymbol{u})\boldsymbol{v}, && \forall \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega),\\
&b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})=(B(\boldsymbol{u},\boldsymbol{v}),\boldsymbol{w})=\frac{1}{2}\int_\Omega(\boldsymbol{u}\cdot\nabla\boldsymbol{v})\cdot\boldsymbol{w}dx-\frac{1}{2}\int_\Omega(\boldsymbol{u}\cdot\nabla\boldsymbol{w})\cdot\boldsymbol{v}dx, && \forall \boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\boldsymbol{H}_0^1(\Omega).
\end{aligned}$$

For trilinear form $b(\cdot,\cdot,\cdot)$, it is not hard to derive following inequalities from Sobolev inequalities and Hölder inequality, see e.g., [12, 13]:

$$|b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})|\leq C\|\boldsymbol{u}\|_1\|\boldsymbol{v}\|_1\|\boldsymbol{w}\|_1, \tag{2.1a}$$

$$|b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})|\leq C\|\boldsymbol{u}\|_1\|\boldsymbol{v}\|_2\|\boldsymbol{w}\|, \tag{2.1b}$$

$$|b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})|\leq C\|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_1\|\boldsymbol{w}\|. \tag{2.1c}$$

## 2.2 Nonconforming $P_1$ finite element

We briefly review the nonconforming $P_1$ finite element space proposed by Crouzeix and Raviart [9]. Let $\mathscr{T}_h$ be a shape-regular triangulation in $\Omega$. Let $h$ be the maximal diameter over all elements in $\mathscr{T}_h$. The nonconforming $P_1$ finite element space $\boldsymbol{V}_h$ is defined as

$$\begin{aligned}\boldsymbol{V}_h=\{&\boldsymbol{v}\in\boldsymbol{L}^2(\Omega)|\boldsymbol{v}|_K\in\boldsymbol{P}_1(K),\ \forall K\in\mathscr{T}_h,\\&\boldsymbol{v}\text{ is continuous at the midpoints of edges of each triangular},\\&\boldsymbol{v}=0\text{ at the midpoints of edges along }\Gamma\}.\end{aligned}$$

It is known that $\boldsymbol{V}_h$ is not a subspace of $\boldsymbol{H}_0^1(\Omega)$. The discrete gradient and divergence operators are given by

$$\nabla_h\boldsymbol{v}|_K=\nabla(\boldsymbol{v}|_K),\quad \mathrm{div}_h\boldsymbol{v}|_K=\mathrm{div}(\boldsymbol{v}|_K),\quad \forall K\in\mathscr{T}_h,$$

for $\boldsymbol{v}\in\boldsymbol{V}_h$. The inverse inequality holds:

$$\|\nabla_h\boldsymbol{v}\|\le\frac{C}{h}\|\boldsymbol{v}\|,\quad \forall\boldsymbol{v}\in\boldsymbol{V}_h. \tag{2.2}$$

We also define the discrete Laplace operator $\Delta_h$:

$$(\Delta_h\boldsymbol{u},\boldsymbol{v})=-(\nabla_h\boldsymbol{u},\nabla_h\boldsymbol{v}),\quad \forall\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{V}_h,$$

and the discrete analogue of $b(\cdot,\cdot,\cdot)$:

$$b_h(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})=\frac{1}{2}\int_\Omega(\boldsymbol{u}\cdot\nabla_h\boldsymbol{v})\cdot\boldsymbol{w}dx-\frac{1}{2}\int_\Omega(\boldsymbol{u}\cdot\nabla_h\boldsymbol{w})\cdot\boldsymbol{v}dx,\quad \boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\boldsymbol{V}_h.$$

The following three terms appear from integration by parts.

$$\begin{aligned}&X_1(\boldsymbol{u},\boldsymbol{v})=\int_\Omega\nabla\boldsymbol{u}:\nabla_h\boldsymbol{v}dx+\int_\Omega\Delta\boldsymbol{u}\cdot\boldsymbol{v}dx, &&\forall\boldsymbol{u}\in\boldsymbol{H}^2(\Omega)\cap\boldsymbol{H}_0^1(\Omega), &&\boldsymbol{v}\in\boldsymbol{H},\\&X_2(u,\boldsymbol{v})=\int_\Omega u\mathrm{div}_h\boldsymbol{v}dx+\int_\Omega\nabla u\cdot\boldsymbol{v}dx, &&\forall u\in H_0^1(\Omega)\cap L_0^2(\Omega), &&\boldsymbol{v}\in\boldsymbol{H},\\&X_3(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})=\frac{1}{2}\sum_{K\in\mathscr{T}_h}\int_{\partial K}(\boldsymbol{u}\cdot\boldsymbol{n})(\boldsymbol{v}\cdot\boldsymbol{w})ds, &&\forall\boldsymbol{u}\in\boldsymbol{H}^2(\Omega)\cap\boldsymbol{H}_0^1(\Omega), &&\boldsymbol{v},\boldsymbol{w}\in\boldsymbol{H},\end{aligned}$$

where $\boldsymbol{H}:=\boldsymbol{H}_0^1(\Omega)+\boldsymbol{V}_h$, $\boldsymbol{n}$ refers to the outward pointing unit normal vector of $\partial K$.

The next lemma gives some useful inequalities for three terms $\{X_i\}_{i=1}^3$. The proofs have been comprehensive discussions in [12, p. 269], and hence omit.

**Lemma 2.1.** *The following statements hold*

$$\begin{aligned}&|X_1(\boldsymbol{u},\boldsymbol{v})|\le Ch\|\boldsymbol{u}\|_2\|\nabla_h\boldsymbol{v}\|,\\&|X_2(u,\boldsymbol{v})|\le Ch\|u\|_1\|\nabla_h\boldsymbol{v}\|,\\&|X_3(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})|\le Ch\|\boldsymbol{u}\|_2\|\nabla_h\boldsymbol{v}\|\|\nabla_h\boldsymbol{w}\|.\end{aligned}$$

It is easy to check that

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

Next we introduce the interpolation operator $\iota_h:\boldsymbol{H}_0^1(\Omega)\to\boldsymbol{V}_h$

$$\iota_h\boldsymbol{v}(m_e)=\frac{1}{|e|}\int_e\boldsymbol{v}ds,$$

where $m_e$ represents the midpoint of edge $e$. We have following estimates:

**Lemma 2.2.** *The following statements hold*

$$\|\boldsymbol{u}-\iota_h\boldsymbol{u}\|+h\|\nabla_h(\boldsymbol{u}-\iota_h\boldsymbol{u})\|\le Ch^2\|\boldsymbol{u}\|_2, \tag{2.3a}$$

$$\|\mathrm{div}_h(\boldsymbol{u}-\iota_h\boldsymbol{u})\|\le Ch\|\mathrm{div}\boldsymbol{u}\|_1. \tag{2.3b}$$

*Proof.* The first inequality is proved in [9] and the second is proved in [12, Lemma 3.1]. □

# 3 Assumptions and priori estimates

In this section we give some assumptions related to the external force $\boldsymbol{f}_{ext}(\boldsymbol{x},t)$, initial velocity $\boldsymbol{u}_0(\boldsymbol{x})$ and initial guess of pressure $p_0(\boldsymbol{x},t)$. We will assume that (A1)-(A4) hold true in the remaining of this paper without explicitly noticing.

(A1) $\mathrm{div}\boldsymbol{u}_0=0$, $\boldsymbol{u}_0\in\boldsymbol{H}^2(\Omega)\cap\boldsymbol{H}_0^1(\Omega)$, $\Delta\boldsymbol{u}_0\in\boldsymbol{H}^1(\Omega)$.

(A2) The following Neumann problem is consistent and the initial pressure $p_0(\boldsymbol{x},0)$ to be the solution of it:

$$\begin{cases}\Delta q=\mathrm{div}\big(\boldsymbol{f}_{ext}(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)\big) & \text{in } \Omega,\\ \nabla q=\boldsymbol{f}_{ext}(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)+\Delta\boldsymbol{u}_0 & \text{on } \Gamma.\end{cases}$$

(A3) $\boldsymbol{f}_{ext}\in L^\infty(\boldsymbol{L^2})$, $(\boldsymbol{f}_{ext})_t\in L^2(\boldsymbol{L^2})$, $\nabla p_0\in L^\infty(\boldsymbol{L^2})$, $\nabla(p_0)_t\in L^2(\boldsymbol{L^2})$, i.e.,

$$\begin{aligned}&\sup_{0\le t\le T}\|\boldsymbol{f}_{ext}(\cdot,t)\|+\int_0^T\|(\boldsymbol{f}_{ext})_t(\cdot,t)\|^2dt\le C,\\&\sup_{0\le t\le T}\|\nabla p_0(\cdot,t)\|+\int_0^T\|\nabla(p_0)_t(\cdot,t)\|^2dt\le C.\end{aligned}$$

(A4) $\boldsymbol{f}_{ext}(\boldsymbol{x},0)-\nabla p_0(\boldsymbol{x},0)\in\boldsymbol{H}^1(\Omega)$.

**Remark 3.1.** The over-determined Neumann problem in (A2) is studied in [10] and several sufficient conditions for the existence of solution are given.

**Lemma 3.1.** *For* $s=1,2,\ldots$, *we have*

$$\boldsymbol{f}_s(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)\in\boldsymbol{H}^1(\Omega),\quad \operatorname{div}(\boldsymbol{f}_s(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0))=0, \tag{i}$$

$$\sup_{0\le t\le T}\|\boldsymbol{f}_s(\cdot,t)\|+\int_0^T\|(\boldsymbol{f}_s)_t(\cdot,t)\|^2dt\le C, \tag{ii}$$

$$\sup_{0\le t\le T}\|\nabla p_s(\cdot,t)\|+\int_0^T\|(\nabla p_s)_t(\cdot,t)\|^2dt\le C, \tag{iii}$$

$$p_s(\boldsymbol{x},0)=p_0(\boldsymbol{x},0). \tag{iv}$$

*Proof.* It suffices to show the case $s=1$. By definition and Assumption (A1), we have

$$B(\boldsymbol{u}_0,\boldsymbol{u}_0)=\boldsymbol{u}_0\cdot\nabla\boldsymbol{u}_0\in\boldsymbol{H}^1(\Omega).$$

Since $\boldsymbol{f}_1=\boldsymbol{f}_{ext}-\nabla p_0$, Assumptions (A4) and (A2) imply that

$$\boldsymbol{f}_1(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)=\boldsymbol{f}_{ext}(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)-\nabla p_0(\boldsymbol{x},0)\in\boldsymbol{H}^1(\Omega), \tag{3.1a}$$

$$\operatorname{div}\big(\boldsymbol{f}_1(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)\big)=\operatorname{div}\big(\boldsymbol{f}_{ext}(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)\big)-\Delta p_0(\boldsymbol{x},0)=0. \tag{3.1b}$$

This proves the assertion (i). By triangle inequality and Assumption (A3), we have

$$\sup_{0\le t\le T}\|\boldsymbol{f}_1(\cdot,t)\|\le\sup_{0\le t\le T}\|\boldsymbol{f}_{ext}(\cdot,t)\|+\sup_{0\le t\le T}\|\nabla p_0(\cdot,t)\|\le C, \tag{3.2a}$$

$$\int_0^T\|(\boldsymbol{f}_1)_t(\cdot,t)\|^2dt\le 2\left(\int_0^T\|(\boldsymbol{f}_{ext})_t(\cdot,t)\|^2+\|\nabla(p_0)_t(\cdot,t)\|^2dt\right)\le C. \tag{3.2b}$$

Next we turn to show (iii). Using the identity (1.2b), we have

$$p_1=p_0-\frac{1}{\epsilon}\operatorname{div}\big((\boldsymbol{u}_1)_t+\alpha\boldsymbol{u}_1\big).$$

Then the priori estimate (3.2), Assumption (A1) and [13, Lemma 4.2] lead to

$$\frac{1}{\epsilon}\nabla\operatorname{div}\big((\boldsymbol{u}_1)_t+\alpha\boldsymbol{u}_1\big)\in L^\infty(\boldsymbol{L}^2).$$

Let

$$\boldsymbol{g}:=\boldsymbol{f}_1(\boldsymbol{x},0)+\Delta\boldsymbol{u}_0-B(\boldsymbol{u}_0,\boldsymbol{u}_0).$$

Then by the properties (3.1) and Assumption (A1), we derive that $\boldsymbol{g}\in\boldsymbol{H}^1(\Omega)$ and $\text{div}\boldsymbol{g}=0$. Consequently, we have [13, Lemma 4.3]

$$\frac{1}{\epsilon}\nabla\text{div}\big((\boldsymbol{u}_1)_{tt}+\alpha(\boldsymbol{u}_1)_t\big)\in L^2(\boldsymbol{L}^2).$$

These two priori regularities and Assumption (A3) show that $\nabla p_1\in L^\infty(\boldsymbol{L}^2)$ and $\nabla(p_1)_t\in L^2(\boldsymbol{L}^2)$. Now it remains to prove the assertion (iv). Let $\boldsymbol{h}=(\boldsymbol{u}_1)_t(\cdot,0)$. Then by formulation (1.3a), it satisfies

$$\boldsymbol{h}-\frac{1}{\epsilon}\nabla\text{div}\boldsymbol{h}=\boldsymbol{g}\quad\text{in}\ \ \Omega. \tag{3.3}$$

Now let $w\in C_0^\infty(\Omega)$ solve

$$w-\frac{1}{\epsilon}\Delta w=v\qquad\text{for any}\ \ v\in C_0^\infty(\Omega),$$

then direct computation give

$$\begin{aligned}(\text{div}\boldsymbol{h},v)&=\left(\text{div}\boldsymbol{h},w-\frac{1}{\epsilon}\Delta w\right)=-(\boldsymbol{h},\nabla w)+\left(\frac{1}{\epsilon}\nabla\text{div}\boldsymbol{h},\nabla w\right)\\&=(\text{div}\boldsymbol{g},w)=0,\quad\forall v\in C_0^\infty(\Omega),\end{aligned}$$

where the last step follows from the fact $\text{div}\boldsymbol{g}=0$. This implies that $\text{div}\boldsymbol{h}=0$ in the distribution sense, and hence $p_1(\boldsymbol{x},0)=p_0(\boldsymbol{x},0)$ follows from (1.2b) immediately.

Now for any integer $s\geq 2$, we only sketch the proof by mathematics induction. Assume that the assertions hold for $s=k-1$, and we show they also hold for $s=k$. Firstly by inductive hypothesis and Assumption (A4), there holds

$$\boldsymbol{f}_k(\boldsymbol{x},0)=\boldsymbol{f}_{ext}(\boldsymbol{x},0)-\nabla p_{k-1}(\boldsymbol{x},0)=\boldsymbol{f}_{ext}(\boldsymbol{x},0)-\nabla p_0(\boldsymbol{x},0)\in\boldsymbol{H}^1(\Omega),$$

and

$$\text{div}\big(\boldsymbol{f}_k(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)\big)=0$$

follows from Assumption (A2) immediately. Next by applying triangle inequality, we obtain $\boldsymbol{f}_k\in L^\infty(\boldsymbol{L}^2)$ and $(\boldsymbol{f}_k)_t\in L^2(\boldsymbol{L}^2)$. Now define by

$$\widetilde{\boldsymbol{g}}=\boldsymbol{f}_k(\boldsymbol{x},0)+\Delta\boldsymbol{u}_0-B(\boldsymbol{u}_0,\boldsymbol{u}_0),$$

Assumption (A1) and direct computation show that $\widetilde{\boldsymbol{g}}\in\boldsymbol{H}^1(\Omega)$ with $\text{div}\widetilde{\boldsymbol{g}}=0$. Then identical argument derives that $\nabla p_k\in L^\infty(\boldsymbol{L}^2)$ and $\nabla(p_k)_t\in L^2(\boldsymbol{L}^2)$. Now let $\widetilde{\boldsymbol{h}}=(\boldsymbol{u}_k)_t(\cdot,0)$, then it also satisfies

$$\widetilde{\boldsymbol{h}}-\frac{1}{\epsilon}\nabla\text{div}\widetilde{\boldsymbol{h}}=\widetilde{\boldsymbol{g}}\qquad\text{with}\quad\text{div}\widetilde{\boldsymbol{g}}=0.$$

Hence we obtain div$\widetilde{\boldsymbol{h}}=0$ like the case $s=1$, and inductive hypothesis proves that

$$p_k(\boldsymbol{x},0)=p_{k-1}(\boldsymbol{x},0)=p_0(\boldsymbol{x},0).$$

The proof is completed. □

Now we introduce $\boldsymbol{u}_s^{-1}$ $(s=1,2,\ldots)$ by the equation below:

$$\frac{\boldsymbol{u}_s^0-\boldsymbol{u}_s^{-1}}{\Delta t}-\frac{1}{\epsilon}\nabla\mathrm{div}\frac{\boldsymbol{u}_s^0-\boldsymbol{u}_s^{-1}}{\Delta t}-\Delta\boldsymbol{u}_s^0+B(\boldsymbol{u}_s^0,\boldsymbol{u}_s^0)=\boldsymbol{f}_s^0. \tag{3.4}$$

Clearly, the solution $\boldsymbol{u}_s^{-1}$ is well-defined. Moreover we have the following Lemma.

**Lemma 3.2.** *The following properties are true*

$$\mathrm{div}\frac{\boldsymbol{u}_s^0-\boldsymbol{u}_s^{-1}}{\Delta t}=0,\quad \frac{\boldsymbol{u}_s^0-\boldsymbol{u}_s^{-1}}{\Delta t}\in\boldsymbol{H}^1(\Omega),\quad s=1,2,\ldots.$$

*Proof.* Let

$$\boldsymbol{h}=\frac{\boldsymbol{u}_s^0-\boldsymbol{u}_s^{-1}}{\Delta t}.$$

By Eq. (3.4) and Lemma 3.1, we have

$$\begin{aligned}\boldsymbol{h}-\frac{1}{\epsilon}\nabla\mathrm{div}\boldsymbol{h}&=\boldsymbol{f}_s^0-B(\boldsymbol{u}_s^0,\boldsymbol{u}_s^0)+\Delta\boldsymbol{u}_s^0\\&=\boldsymbol{f}_{ext}(\boldsymbol{x},0)-\nabla p_0(\boldsymbol{x},0)-B(\boldsymbol{u}_0,\boldsymbol{u}_0)+\Delta\boldsymbol{u}_0=:\boldsymbol{g},\end{aligned} \tag{3.5}$$

where we have used

$$\boldsymbol{f}_s^0=\boldsymbol{f}_s(\boldsymbol{x},0)=\boldsymbol{f}_{ext}(\boldsymbol{x},0)-\nabla p_{s-1}(\boldsymbol{x},0)=\boldsymbol{f}_{ext}(\boldsymbol{x},0)-\nabla p_0(\boldsymbol{x},0).$$

As the proof in Lemma 3.1, direct computation gives $\boldsymbol{g}\in\boldsymbol{H}^1(\Omega)$ and div$\boldsymbol{g}=0$. Consequently, we also have div$\boldsymbol{h}=0$ by the identical argument in [13, Lemma 4.3]. Using Eq. (3.5), we have $\boldsymbol{h}=\boldsymbol{g}\in\boldsymbol{H}^1(\Omega)$. The proof is completed. □

In the sequel, we omit index $s$ of $\boldsymbol{u}$ and $\boldsymbol{f}$ for simplicity. Consider semi-discretization scheme: find $\{\boldsymbol{u}^n\}_{n=1}^N$ with $\boldsymbol{u}^0=\boldsymbol{u}_0$ and the zero Dirichlet boundary condition, such that

$$\begin{aligned}&\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}-\frac{1}{\epsilon}\nabla\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}+\alpha\mathrm{div}\boldsymbol{u}^n\right)-\Delta\boldsymbol{u}^n\\&\quad+B(\boldsymbol{u}^{n-1},\boldsymbol{u}^n)=\boldsymbol{f}^n,\quad n=1,2,\cdots,N,\end{aligned} \tag{3.6}$$

and full discretization scheme: find $\{\boldsymbol{u}_h^n\}_{n=1}^N \subset \boldsymbol{V}_h$ with $\boldsymbol{u}_h^0 = \iota_h \boldsymbol{u}_0$ such that

$$\left(\frac{\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1}}{\Delta t}, \boldsymbol{v}_h\right) + \frac{1}{\epsilon}\left(\mathrm{div}_h \frac{\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1}}{\Delta t} + \alpha \mathrm{div}_h \boldsymbol{u}_h^n, \mathrm{div}_h \boldsymbol{v}_h\right)$$
$$+ (\nabla_h \boldsymbol{u}_h^n, \nabla_h \boldsymbol{v}_h) + b_h(\boldsymbol{u}_h^{n-1}, \boldsymbol{u}_h^n, \boldsymbol{v}_h) = (\boldsymbol{f}^n, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{3.7}$$

for every $n = 1, 2, \cdots, N$.

From Assumption (A2) and Lemma 3.1, we have

$$\mathrm{div}\big(\boldsymbol{f}^0 - B(\boldsymbol{u}_0, \boldsymbol{u}_0)\big) = 0, \quad \boldsymbol{f}^0 \in \boldsymbol{H}^1(\Omega),$$
$$\sup_{0 \le t \le T} \|\boldsymbol{f}(\cdot,t)\| + \int_0^T \|\boldsymbol{f}_t(\cdot,t)\|^2 dt \le C.$$

From Lemma 3.2, we have

$$\mathrm{div} \frac{\boldsymbol{u}^0 - \boldsymbol{u}^{-1}}{\Delta t} = 0, \quad \frac{\boldsymbol{u}^0 - \boldsymbol{u}^{-1}}{\Delta t} \in \boldsymbol{H}^1(\Omega). \tag{3.8}$$

These relations will be frequently used through our analysis.

**Lemma 3.3.** *Let $\{\boldsymbol{u}^n\}_{n=0}^N$ and $\boldsymbol{u}^{-1}$ be the solution of the semi-discretization scheme* (3.6) *and Eq.* (3.4)*, respectively. Then for sufficiently small $\epsilon$, we have*

$$\frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{u}^n - \boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2 + \frac{1}{\epsilon^2}\|\mathrm{div}\boldsymbol{u}^n\|_1^2 + \|\boldsymbol{u}^n\|_2^2 \le C, \quad n = 0, 1, \cdots, N.$$

*Proof.* All the estimates are already contained in [13, Lemma 3.1 and Lemma 4.5], and the range of $\epsilon$ is given in [13, Lemma 3.1]. □

**Lemma 3.4.** *Let $\{\boldsymbol{u}^n\}_{n=0}^N$ and $\boldsymbol{u}^{-1}$ be the solution of the semi-discretization scheme* (3.6) *and Eq.* (3.4)*, respectively. Then for sufficiently small $\epsilon$, we have*

$$\sum_{n=1}^N \left(\frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{u}^n - 2\boldsymbol{u}^{n-1} + \boldsymbol{u}^{n-2}}{\Delta t^2}\right\|_1^2 + \frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{u}^n - \boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2 + \left\|\frac{\boldsymbol{u}^n - \boldsymbol{u}^{n-1}}{\Delta t}\right\|_2^2\right)\Delta t \le C.$$

*Proof.* The Lemma is a discrete version of [13, Lemma 4.3]. Define by

$$A\boldsymbol{u}^n := -\frac{1}{\epsilon}\nabla\left(\mathrm{div}\frac{\boldsymbol{u}^n - \boldsymbol{u}^{n-1}}{\Delta t} + \alpha \mathrm{div}\boldsymbol{u}^n\right) - \Delta \boldsymbol{u}^n \quad \text{for} \quad n = 0, 1, \cdots, N.$$

Take the backward difference quotient in (1.4a), multiply $A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}$ and integrate over the domain $\Omega$ on both sides lead to

$$\begin{aligned}&\left(\frac{\boldsymbol{u}^n-2\boldsymbol{u}^{n-1}+\boldsymbol{u}^{n-2}}{\Delta t^2},A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right)+\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2\\ =&\left(\frac{\boldsymbol{f}^n-\boldsymbol{f}^{n-1}}{\Delta t},A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right)-b\left(\boldsymbol{u}^{n-1},\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right)\\ &-b\left(\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t},\boldsymbol{u}^{n-1},A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right):=\mathrm{I}_1+\mathrm{I}_2+\mathrm{I}_3. \end{aligned}\tag{3.9}$$

Then by Young's inequality, we have

$$|\mathrm{I}_1|\le\left\|\frac{\boldsymbol{f}^n-\boldsymbol{f}^{n-1}}{\Delta t}\right\|^2+\frac{1}{4}\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2.$$

Next using Young's inequality and the properties (2.1) of trilinear form $b(\cdot,\cdot,\cdot)$, we also have

$$\begin{aligned}|\mathrm{I}_2|&\le C\|\boldsymbol{u}^{n-1}\|_2\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_1\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|\\ &\le C\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2+\frac{1}{4}\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2,\\ |\mathrm{I}_3|&\le C\left\|\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t}\right\|_1\|\boldsymbol{u}^{n-1}\|_2\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|\\ &\le C\left\|\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t}\right\|_1^2+\frac{1}{4}\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2,\end{aligned}$$

where the last step is due to Lemma 3.3. Now by definition of the operator $A$ and direct computation, we have the following splitting

$$\begin{aligned}&\left(\frac{\boldsymbol{u}^n-2\boldsymbol{u}^{n-1}+\boldsymbol{u}^{n-2}}{\Delta t^2},A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right)\\ =&\frac{1}{\epsilon}\left\|\mathrm{div}\frac{\boldsymbol{u}^n-2\boldsymbol{u}^{n-1}+\boldsymbol{u}^{n-2}}{\Delta t}\right\|^2+\frac{\alpha}{2\epsilon\Delta t}\left(\left\|\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2-\left\|\mathrm{div}\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t}\right\|^2\right.\\ &\left.+\left\|\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}-\mathrm{div}\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t}\right\|^2\right)+\frac{1}{2\Delta t}\left(\left\|\nabla\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2\right.\\ &\left.-\left\|\nabla\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t}\right\|^2+\left\|\nabla\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}-\nabla\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t}\right\|^2\right).\end{aligned}$$

Multiplying $\Delta t$ on both sides of equation (3.9) and summing it over $n$ from 1 to $k\in\{1,2,\cdots,N\}$ and using the property (3.8), together with three estimates above, we obtain

$$\begin{aligned}&\frac{1}{\epsilon}\left\|\mathrm{div}\frac{\boldsymbol{u}^k-\boldsymbol{u}^{k-1}}{\Delta t}\right\|^2+\left\|\nabla\frac{\boldsymbol{u}^k-\boldsymbol{u}^{k-1}}{\Delta t}\right\|^2+\sum_{n=1}^{k}\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2\Delta t\\ \le& C+C\sum_{n=1}^{k}\left\|\frac{\boldsymbol{f}^n-\boldsymbol{f}^{n-1}}{\Delta t}\right\|^2\Delta t+C\sum_{n=0}^{k}\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2\Delta t\\ \le& C+\sum_{n=0}^{k}\left\|\nabla\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2\Delta t,\end{aligned}$$

where the last line is due to Poincáre inequality and the elementary estimate

$$\sum_{n=1}^{k}\|\frac{\boldsymbol{f}^n-\boldsymbol{f}^{n-1}}{\Delta t}\|^2\Delta t\le\int_0^T\|\boldsymbol{f}_t\|^2dt\le C,$$

see [13, p. 1490]. This and discrete Gronwall's inequality imply that for sufficiently small $\Delta t$, there holds

$$\sum_{n=1}^{N}\left\|A\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|^2\Delta t\le C.$$

We complete the proof by [13, Lemma 3.1]. □

# 4 Error estimates to the full discretization scheme

To get the error estimation, we first split error into the linear part and nonlinear part, by introducing an auxiliary problem: find $\{\boldsymbol{u}_*^n\}_{n=0}^N\subset\boldsymbol{V}_h$ with $\boldsymbol{u}_*^0=\iota_h\boldsymbol{u}_0$ such that

$$\begin{aligned}&\left(\frac{\boldsymbol{u}_*^n-\boldsymbol{u}_*^{n-1}}{\Delta t},\boldsymbol{v}_h\right)+\frac{1}{\epsilon}\left(\mathrm{div}_h\frac{\boldsymbol{u}_*^n-\boldsymbol{u}_*^{n-1}}{\Delta t},\mathrm{div}_h\boldsymbol{v}_h\right)+\frac{\alpha}{\epsilon}(\mathrm{div}_h\boldsymbol{u}_*^n,\mathrm{div}_h\boldsymbol{v}_h)\\ &\quad+(\nabla_h\boldsymbol{u}_*^n,\nabla_h\boldsymbol{v}_h)+b_h(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\boldsymbol{v}_h)=(\boldsymbol{f}^n,\boldsymbol{v}_h),\quad\forall\boldsymbol{v}_h\in\boldsymbol{V}_h,\quad n=1,2,\cdots,N.\end{aligned}\tag{4.1}$$

Next we define by $\boldsymbol{e}^n:=\boldsymbol{u}^n-\boldsymbol{u}_h^n=\boldsymbol{e}_1^n+\boldsymbol{e}_2^n$ for $n=0,1,\cdots,N$, where $\boldsymbol{e}_1^n:=\boldsymbol{u}^n-\boldsymbol{u}_*^n$ and $\boldsymbol{e}_2^n:=\boldsymbol{u}_*^n-\boldsymbol{u}_h^n$. Then by formulations (4.1) and (3.6), we know that $\{\boldsymbol{e}_1^n\}_{n=0}^N$ with

$\boldsymbol{e}_1^0=\boldsymbol{u}_0-\iota_h\boldsymbol{u}_0$ satisfies

$$\begin{aligned}&\left(\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\boldsymbol{v}_h\right)+\frac{1}{\epsilon}\left(\mathrm{div}_h\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\mathrm{div}_h\boldsymbol{v}_h\right)+\frac{\alpha}{\epsilon}(\mathrm{div}_h\boldsymbol{e}_1^n,\mathrm{div}_h\boldsymbol{v}_h)+(\nabla_h\boldsymbol{e}_1^n,\nabla_h\boldsymbol{v}_h)\\&=X_1(\boldsymbol{u}^n,\boldsymbol{v}_h)+\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{v}_h\right)+\frac{\alpha}{\epsilon}X_2(\mathrm{div}\boldsymbol{u}^n,\boldsymbol{v}_h)\\&\quad-X_3(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\boldsymbol{v}_h),\quad\forall\boldsymbol{v}_h\in\boldsymbol{V}_h,\quad n=1,2,\cdots,N.\end{aligned}\tag{4.2}$$

Moreover, by formulations (3.7) and (4.1), we also know that $\{\boldsymbol{e}_2^n\}_{n=0}^N\subset\boldsymbol{V}_h$ with $\boldsymbol{e}_2^0=0$ satisfies

$$\begin{aligned}&\left(\frac{\boldsymbol{e}_2^n-\boldsymbol{e}_2^{n-1}}{\Delta t},\boldsymbol{v}_h\right)+\frac{1}{\epsilon}\left(\mathrm{div}_h\frac{\boldsymbol{e}_2^n-\boldsymbol{e}_2^{n-1}}{\Delta t},\mathrm{div}_h\boldsymbol{v}_h)+\frac{\alpha}{\epsilon}(\mathrm{div}_h\boldsymbol{e}_2^n,\mathrm{div}_h\boldsymbol{v}_h\right)+(\nabla_h\boldsymbol{e}_2^n,\nabla_h\boldsymbol{v}_h)\\&\quad+b_h(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\boldsymbol{v}_h)-b_h(\boldsymbol{u}_h^{n-1},\boldsymbol{u}_h^n,\boldsymbol{v}_h)=0,\quad\forall\boldsymbol{v}_h\in\boldsymbol{V}_h,\quad n=1,2,\cdots,N.\end{aligned}\tag{4.3}$$

The next two lemmas provides the error bounds of the term $\boldsymbol{e}_1^n$ in $\boldsymbol{H}^1$ and $\boldsymbol{L}^2$, respectively.

**Lemma 4.1.** *Assume that $\epsilon$ is small enough. Then we have*

$$\begin{aligned}&\sum_{n=1}^N\left\|\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t}\right\|^2\Delta t+\frac{1}{\epsilon}\sum_{n=1}^N\left\|\mathrm{div}_h\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t}\right\|^2\Delta t\\&\quad+\max_{n=1,\cdots,N}\frac{1}{\epsilon}\|\mathrm{div}_h\boldsymbol{e}_1^n\|^2+\max_{n=1,\cdots,N}\|\nabla_h\boldsymbol{e}_1^n\|^2\le Ch^2.\end{aligned}$$

*Proof.* Let $\boldsymbol{z}_1^n=\boldsymbol{u}^n-\iota_h\boldsymbol{u}^n$, $\boldsymbol{z}_2^n=\iota_h\boldsymbol{u}^n-\boldsymbol{u}_*^n$, then $\boldsymbol{e}_1^n=\boldsymbol{z}_1^n+\boldsymbol{z}_2^n$. Inputting

$$\boldsymbol{v}_h=\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\in\boldsymbol{V}_h$$

into the weak formulation (4.2), we have

$$\begin{aligned}&\left\|\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2+\frac{1}{\epsilon}\left\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\\&\quad+\frac{\alpha}{\epsilon}\left(\mathrm{div}_h\boldsymbol{z}_2^n,\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)+\left(\nabla_h\boldsymbol{z}_2^n,\nabla_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)\\&=-\left(\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t},\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)-\frac{1}{\epsilon}\left(\mathrm{div}_h\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t},\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)\\&\quad-\frac{\alpha}{\epsilon}\left(\mathrm{div}_h\boldsymbol{z}_1^n,\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)-\left(\nabla_h\boldsymbol{z}_1^n,\nabla_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)+X_1\left(\boldsymbol{u}^n,\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)\end{aligned}$$

$$+\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)+\frac{\alpha}{\epsilon}X_2\left(\mathrm{div}\boldsymbol{u}^n,\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)$$
$$-X_3\left(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right)=:\sum_{i=1}^{8}\mathrm{I}_i^n. \tag{4.4}$$

Now we bound these terms separately. By Young's inequality and the approximation property (2.3) of $\iota_h$, there holds

$$\begin{aligned}
\mathrm{I}_1^n&\le\frac{1}{2}\left\|\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t}\right\|^2+\frac{1}{2}\left\|\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\le\frac{1}{2}\left\|\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2+Ch^2\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_2^2,\\
\mathrm{I}_2^n&\le\frac{1}{\epsilon}\left\|\mathrm{div}_h\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t}\right\|^2+\frac{1}{4\epsilon}\left\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\\
&\le\frac{1}{4\epsilon}\left\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2+C\frac{h^2}{\epsilon}\left\|\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2,\\
\mathrm{I}_3^n&\le\frac{C}{\epsilon}\left\|\mathrm{div}_h\boldsymbol{z}_1^n\right\|^2+\frac{1}{4\epsilon}\left\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\le\frac{1}{4\epsilon}\left\|\mathrm{div}_h\frac{\boldsymbol{z}_2^{n+1}-\boldsymbol{z}_2^n}{\Delta t}\right\|^2+C\frac{h^2}{\epsilon}\|\mathrm{div}\boldsymbol{u}^n\|_1^2.
\end{aligned}$$

Next we estimate the terms $\mathrm{I}_4$ and $\mathrm{I}_5$. Using the summation by parts formula, the approximation property (2.3) of $\iota_h$ and Young's inequality, we arrive at

$$\begin{aligned}
\mathrm{I}_4^n&=-\frac{1}{\Delta t}\left[(\nabla_h\boldsymbol{z}_1^n,\nabla_h\boldsymbol{z}_2^n)-(\nabla_h\boldsymbol{z}_1^{n-1},\nabla_h\boldsymbol{z}_2^{n-1})\right]+\left(\nabla_h\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t},\nabla_h\boldsymbol{z}_2^{n-1}\right)\\
&\le-\frac{1}{\Delta t}\left[(\nabla_h\boldsymbol{z}_1^n,\nabla_h\boldsymbol{z}_2^n)-(\nabla_h\boldsymbol{z}_1^{n-1},\nabla_h\boldsymbol{z}_2^{n-1})\right]+C\left\|\nabla_h\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t}\right\|^2+C\|\nabla_h\boldsymbol{z}_2^{n-1}\|^2\\
&\le-\frac{1}{\Delta t}\left[(\nabla_h\boldsymbol{z}_1^n,\nabla_h\boldsymbol{z}_2^n)-(\nabla_h\boldsymbol{z}_1^{n-1},\nabla_h\boldsymbol{z}_2^{n-1})\right]+Ch^2\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_2^2+C\|\nabla_h\boldsymbol{z}_2^{n-1}\|^2,
\end{aligned}$$

and together with Lemma 2.1, we also have

$$\begin{aligned}
\mathrm{I}_5^n&=\frac{1}{\Delta t}\left[X_1(\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_1(\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]-X_1\left(\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{z}_2^{n-1}\right)\\
&\le\frac{1}{\Delta t}\left[X_1(\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_1(\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]+Ch\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_2\|\nabla_h\boldsymbol{z}_2^{n-1}\|\\
&\le\frac{1}{\Delta t}\left[X_1(\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_1(\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]+Ch^2\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_2^2+C\|\nabla_h\boldsymbol{z}_2^{n-1}\|^2.
\end{aligned}$$

Similarly, there holds for the terms $\mathrm{I}_6$ and $\mathrm{I}_7$

$$\begin{aligned}
\mathrm{I}_6^n=&\frac{1}{\epsilon\Delta t}\left[X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{z}_2^n\right)-X_2\left(\mathrm{div}\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t},\boldsymbol{z}_2^{n-1}\right)\right]\\
&-\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-2\boldsymbol{u}^{n-1}+\boldsymbol{u}^{n-2}}{\Delta t^2},\boldsymbol{z}_2^{n-1}\right)\\
\le&\frac{1}{\epsilon\Delta t}\left[X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{z}_2^n\right)-X_2\left(\mathrm{div}\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t},\boldsymbol{z}_2^{n-1}\right)\right]\\
&+C\frac{h}{\epsilon}\left\|\mathrm{div}\frac{\boldsymbol{u}^n-2\boldsymbol{u}^{n-1}+\boldsymbol{u}^{n-2}}{\Delta t^2}\right\|_1\|\nabla_h\boldsymbol{z}_2^{n-1}\|\\
\le&\frac{1}{\epsilon\Delta t}\left[X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{z}_2^n\right)-X_2\left(\mathrm{div}\frac{\boldsymbol{u}^{n-1}-\boldsymbol{u}^{n-2}}{\Delta t},\boldsymbol{z}_2^{n-1}\right)\right]\\
&+C\frac{h^2}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{u}^n-2\boldsymbol{u}^{n-1}+\boldsymbol{u}^{n-2}}{\Delta t^2}\right\|_1^2+C\|\nabla_h\boldsymbol{z}_2^{n-1}\|^2,\\
\mathrm{I}_7^n=&\frac{\alpha}{\epsilon\Delta t}\left[X_2(\mathrm{div}\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_2(\mathrm{div}\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]-\frac{\alpha}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{z}_2^{n-1}\right)\\
\le&\frac{\alpha}{\epsilon\Delta t}\left[X_2(\mathrm{div}\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_2(\mathrm{div}\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]+C\frac{h}{\epsilon}\left\|\mathrm{div}\frac{\boldsymbol{u}^{n+1}-\boldsymbol{u}^n}{\Delta t}\right\|_1\|\nabla_h\boldsymbol{z}_2^{n-1}\|\\
\le&\frac{\alpha}{\epsilon\Delta t}\left[X_2(\mathrm{div}\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_2(\mathrm{div}\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]+C\frac{h^2}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{u}^{n+1}-\boldsymbol{u}^n}{\Delta t}\right\|_1^2+C\|\nabla_h\boldsymbol{z}_2^{n-1}\|^2.
\end{aligned}$$

For the term $\mathrm{I}_8$, the following splitting, Lemma 2.1 and Young's inequality imply

$$\begin{aligned}
\mathrm{I}_8^n=&-\frac{1}{\Delta t}\left[X_3(\boldsymbol{u}^n,\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_3(\boldsymbol{u}^{n-1},\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]\\
&+X_3\left(\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{u}^n,\boldsymbol{z}_2^n\right)+X_3\left(\boldsymbol{u}^{n-1},\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\boldsymbol{z}_2^{n-1}\right)\\
\le&-\frac{1}{\Delta t}\left[X_3(\boldsymbol{u}^n,\boldsymbol{u}^n,\boldsymbol{z}_2^n)-X_3(\boldsymbol{u}^{n-1},\boldsymbol{u}^{n-1},\boldsymbol{z}_2^{n-1})\right]+Ch^2\times\\
&\left(\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_2^2\|\boldsymbol{u}^n\|_1^2+\|\boldsymbol{u}^{n-1}\|_2^2\left\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2\right)+C\left(\|\nabla_h\boldsymbol{z}_2^{n-1}\|+\|\nabla_h\boldsymbol{z}_2^n\|\right).
\end{aligned}$$

Summing up (4.4) over $n$ from 1 to $k\in\{1,2,\cdots,N\}$ and noting $\boldsymbol{z}_2^0=\boldsymbol{0}$, then we get

$$\begin{aligned}
&\sum_{n=1}^k\left\|\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\Delta t+\frac{1}{\epsilon}\sum_{n=1}^k\left\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\Delta t\\
&\quad+\frac{\alpha}{\epsilon}\|\mathrm{div}_h\boldsymbol{z}_2^k\|^2+\|\nabla_h\boldsymbol{z}_2^k\|^2\le C\left(\sum_{n=1}^k\sum_{i=1}^8\mathrm{I}_i^n\right)\Delta t.
\end{aligned}$$

By the preliminary estimates on the terms $\{\mathrm{I}_i^n\}_{i=1}^8$, we have

$$\begin{aligned}\Big(\sum_{n=1}^k\sum_{i=1}^3\mathrm{I}_i^n\Big)\Delta t\le&\frac{1}{2}\sum_{n=1}^k\Big\|\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\Big\|^2\Delta t+\frac{1}{2\epsilon}\sum_{n=1}^k\Big\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\Big\|^2\Delta t\\&+Ch^2\sum_{n=1}^k\Big(\Big\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\Big\|_2^2+\frac{1}{\epsilon}\Big\|\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\Big\|_1^2+\frac{1}{\epsilon}\|\mathrm{div}\boldsymbol{u}^n\|_1^2\Big)\Delta t\\\le&\frac{1}{2}\sum_{n=1}^k\Big\|\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\Big\|^2\Delta t+\frac{1}{2\epsilon}\sum_{n=1}^k\Big\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\Big\|^2\Delta t+Ch^2,\end{aligned}$$

where the last step is due to Lemma 3.3 and 3.4. Using the approximation property (2.3) of $\iota_h$ and noting $\boldsymbol{z}_2^0=\boldsymbol{0}$ imply

$$\begin{aligned}\Big(\sum_{n=1}^k\mathrm{I}_4^n\Big)\Delta t\le&-(\nabla_h\boldsymbol{z}_1^k,\nabla_h\boldsymbol{z}_2^k)+Ch^2\sum_{n=1}^k\Big\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\Big\|_2^2\Delta t+C\sum_{n=0}^k\|\nabla_h\boldsymbol{z}_2^n\|^2\Delta t\\\le&\frac{1}{10}\|\nabla_h\boldsymbol{z}_2^k\|^2+Ch^2\Big(\Big(\sum_{n=1}^k\Big\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\Big\|_2^2\Delta t\Big)+\|\boldsymbol{u}^k\|_2^2\Big)+C\sum_{n=0}^k\|\nabla_h\boldsymbol{z}_2^n\|^2\Delta t\\\le&\frac{1}{10}\|\nabla_h\boldsymbol{z}_2^k\|^2+Ch^2+C\sum_{n=0}^k\|\nabla_h\boldsymbol{z}_2^n\|^2\Delta t,\end{aligned}$$

and similarly, we also have

$$\begin{aligned}\Big(\sum_{n=1}^k\sum_{i=5}^8\mathrm{I}_i^n\Big)\Delta t\le&\frac{4}{10}\|\nabla_h\boldsymbol{z}_2^k\|^2+Ch^2\Big(\sum_{n=1}^k\Big\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\Big\|_2^2\Delta t\\&+\sum_{n=1}^k\frac{1}{\epsilon^2}\Big\|\mathrm{div}\frac{\boldsymbol{u}^n-2\boldsymbol{u}^{n-1}+\boldsymbol{u}^{n-2}}{\Delta t^2}\Big\|_1^2\Delta t+\sum_{n=1}^k\frac{1}{\epsilon^2}\Big\|\frac{\boldsymbol{u}^{n+1}-\boldsymbol{u}^n}{\Delta t}\Big\|_1^2\Delta t\\&+\sum_{n=1}^k\Big(\Big\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\Big\|_2^2\|\boldsymbol{u}^n\|_1^2+\|\boldsymbol{u}^{n-1}\|_2^2\Big\|\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\Big\|_1^2\Big)\Delta t+\|\boldsymbol{u}^k\|_2^2\Big)\\&+C\sum_{n=0}^k\|\nabla_h\boldsymbol{z}_2^n\|^2\Delta t\\\le&\frac{4}{10}\|\nabla_h\boldsymbol{z}_2^k\|^2+Ch^2+C\sum_{n=0}^k\|\nabla_h\boldsymbol{z}_2^n\|^2\Delta t.\end{aligned}$$

This shows that

$$\|\nabla_h\boldsymbol{z}_2^k\|^2\le Ch^2+C\sum_{n=0}^k\|\nabla_h\boldsymbol{z}_2^n\|^2\Delta t$$

for every $k\in\{1,2,\cdots,N\}$. For sufficiently small $\Delta t$, discrete Gronwall's inequality implies $\max_{n=0,1,\cdots N}\|\nabla_h \boldsymbol{z}_2^n\|^2\le Ch^2$. Consequently,

$$\begin{aligned}&\sum_{n=1}^N\left\|\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\Delta t+\frac{1}{\epsilon}\sum_{n=1}^N\left\|\mathrm{div}_h\frac{\boldsymbol{z}_2^n-\boldsymbol{z}_2^{n-1}}{\Delta t}\right\|^2\Delta t\\&\quad+\frac{1}{\epsilon}\max_{n=1,\cdots,N}\|\mathrm{div}_h\boldsymbol{z}_2^n\|^2+\max_{n=1,\cdots,N}\|\nabla_h\boldsymbol{z}_2^n\|^2\le Ch^2.\end{aligned}$$

Meanwhile, by the approximation property (2.3) of $\iota_h$ we know that

$$\begin{aligned}&\sum_{n=1}^N\left\|\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t}\right\|^2\Delta t+\frac{1}{\epsilon}\sum_{n=1}^N\left\|\mathrm{div}_h\frac{\boldsymbol{z}_1^n-\boldsymbol{z}_1^{n-1}}{\Delta t}\right\|^2\Delta t\\&\quad+\frac{1}{\epsilon}\max_{n=1,\cdots,N}\|\mathrm{div}_h\boldsymbol{z}_1^n\|^2+\max_{n=1,\cdots,N}\|\nabla_h\boldsymbol{z}_1^n\|^2\\&\le Ch^2\times\left(\sum_{n=0}^k\left\|\frac{\boldsymbol{u}^{n+1}-\boldsymbol{u}^n}{\Delta t}\right\|_2^2\Delta t+\frac{1}{\epsilon}\sum_{n=0}^k\left\|\mathrm{div}\frac{\boldsymbol{u}^{n+1}-\boldsymbol{u}^n}{\Delta t}\right\|_1^2\Delta t\right.\\&\quad\left.+\frac{1}{\epsilon}\max_{n=1,\cdots,N}\|\mathrm{div}\boldsymbol{u}^n\|_1^2+\max_{n=1,\cdots,N}\|\boldsymbol{u}^n\|_2^2\right)\le Ch^2.\end{aligned}$$

The desired estimate follows from the triangle inequality, and thus the proof is completed. □

To derive $\boldsymbol{L}^2$ error bound on the term $\boldsymbol{e}_1^n$, we introduce the following dual problem: find $\{\boldsymbol{w}^n\}_{n=N}^0$ with $\boldsymbol{w}^N=\boldsymbol{0}$ and the zero Dirichlet boundary condition, such that

$$\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}-\frac{1}{\epsilon}\nabla\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}-\frac{\alpha}{\epsilon}\nabla\mathrm{div}\boldsymbol{w}^{n-1}-\Delta\boldsymbol{w}^{n-1}=\boldsymbol{e}_1^n,\quad n=N,\cdots,1. \tag{4.5}$$

Then we have a crucial priori estimate on the solution $\{\boldsymbol{w}^n\}_{n=N}^0$. The proof is lengthy but standard, and hence is postponed to the appendix.

**Lemma 4.2.** *Let $\{\boldsymbol{w}^n\}_{n=N}^0$ be the solution of problem* (4.5)*. Then there holds*

$$\sum_{n=N}^1\left(\frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right\|_1^2+\frac{1}{\epsilon^2}\|\mathrm{div}\boldsymbol{w}^{n-1}\|_1^2+\|\boldsymbol{w}^{n-1}\|_2^2\right)\Delta t\le C\sum_{n=N}^1\|\boldsymbol{e}_1^n\|^2\Delta t. \tag{4.6}$$

The next lemma provides an estimation for the term $\boldsymbol{e}_1^n$ in $\boldsymbol{L}^2$.

**Lemma 4.3.** *For sufficiently small $\epsilon$, we have*

$$\sum_{n=1}^{N}\|\boldsymbol{e}_1^n\|^2\Delta t\leq Ch^4.$$

*Proof.* We employ the standard duality argument. Multiplying $\boldsymbol{e}_1^n$ and integrating over the domain $\Omega$ on both sides of (4.5), we obtain

$$\begin{aligned}\|\boldsymbol{e}_1^n\|^2=&\left(\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t},\boldsymbol{e}_1^n\right)+\frac{1}{\epsilon}\left(\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t},\mathrm{div}_h\boldsymbol{e}_1^n\right)\\&+\frac{\alpha}{\epsilon}(\mathrm{div}\boldsymbol{w}^{n-1},\mathrm{div}_h\boldsymbol{e}_1^n)+(\nabla\boldsymbol{w}^{n-1},\nabla_h\boldsymbol{e}_1^n)-X_1(\boldsymbol{w}^{n-1},\boldsymbol{e}_1^n)\\&-\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t},\boldsymbol{e}_1^n\right)-\frac{\alpha}{\epsilon}X_2(\mathrm{div}\boldsymbol{w}^{n-1},\boldsymbol{e}_1^n).\end{aligned}\tag{4.7}$$

Meanwhile, inputting $\boldsymbol{v}_h=\iota_h\boldsymbol{w}^{n-1}$ in (4.2) leads to

$$\begin{aligned}&\left(\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\iota_h\boldsymbol{w}^{n-1}\right)+\frac{1}{\epsilon}\left(\mathrm{div}_h\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\mathrm{div}_h\iota_h\boldsymbol{w}^{n-1}\right)\\&\quad+\frac{\alpha}{\epsilon}(\mathrm{div}_h\boldsymbol{e}_1^n,\mathrm{div}_h\iota_h\boldsymbol{w}^{n-1})+(\nabla_h\boldsymbol{e}_1^n,\nabla_h\iota_h\boldsymbol{w}^{n-1})\\&=X_1(\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1})+\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\iota_h\boldsymbol{w}^{n-1}\right)\\&\quad+\frac{\alpha}{\epsilon}X_2(\mathrm{div}\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1})-X_3(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1}).\end{aligned}\tag{4.8}$$

Then by formulations (4.7) and (4.8), we get

$$\|\boldsymbol{e}_1^n\|^2=\mathrm{I}_1^n+\mathrm{I}_2^n+\mathrm{I}_3^n+\mathrm{I}_4^n+\mathrm{I}_5^n,\tag{4.9}$$

where the five terms $\{\mathrm{I}_i^n\}_{i=1}^5$ are given by

$$\begin{aligned}\mathrm{I}_1^n=&\left(\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t},\boldsymbol{e}_1^n\right)-\left(\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\iota_h\boldsymbol{w}^{n-1}\right),\\\mathrm{I}_2^n=&\frac{1}{\epsilon}\left(\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t},\mathrm{div}_h\boldsymbol{e}_1^n\right)-\frac{1}{\epsilon}\left(\mathrm{div}_h\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\mathrm{div}_h\iota_h\boldsymbol{w}^{n-1}\right),\\\mathrm{I}_3^n=&\frac{\alpha}{\epsilon}\left(\mathrm{div}(\boldsymbol{w}^{n-1}-\iota_h\boldsymbol{w}^{n-1}),\mathrm{div}_h\boldsymbol{e}_1^n\right)+\left(\nabla(\boldsymbol{w}^{n-1}-\iota_h\boldsymbol{w}^{n-1}),\nabla_h\boldsymbol{e}_1^n\right),\\\mathrm{I}_4^n=&-X_1\left(\boldsymbol{w}^{n-1},\boldsymbol{e}_1^n\right)-\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t},\boldsymbol{e}_1^n\right)-\frac{\alpha}{\epsilon}X_2\big(\mathrm{div}\boldsymbol{w}^{n-1},\boldsymbol{e}_1^n\big),\\\mathrm{I}_5^n=&X_1\big(\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1}\big)+\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\iota_h\boldsymbol{w}^{n-1}\right)+\frac{\alpha}{\epsilon}X_2\big(\mathrm{div}\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1}\big)\\&-X_3\big(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1}\big).\end{aligned}$$

It suffices to estimate these terms separately. By the summation by parts formula, the approximation property (2.3) of $\iota_h$, Cauchy-Schwarz and Young's inequality, then there holds

$$\begin{aligned}\mathrm{I}_1^n=&\frac{1}{\Delta t}[(\boldsymbol{e}_1^{n-1},\boldsymbol{w}^{n-1})-(\boldsymbol{e}_1^n,\boldsymbol{w}^n)]+\left(\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\boldsymbol{w}^{n-1}-\iota_h\boldsymbol{w}^{n-1}\right)\\ \le&\frac{1}{\Delta t}[(\boldsymbol{e}_1^{n-1},\boldsymbol{w}^{n-1})-(\boldsymbol{e}_1^n,\boldsymbol{w}^n)]+C_\delta h^4\left\|\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t}\right\|^2+\delta\|\boldsymbol{w}^{n-1}\|_2^2,\end{aligned}$$

and similarly,

$$\begin{aligned}\mathrm{I}_2^n=&\frac{1}{\epsilon\Delta t}\left[(\mathrm{div}_h\boldsymbol{e}_1^{n-1},\mathrm{div}\boldsymbol{w}^{n-1})-\left(\mathrm{div}_h\boldsymbol{e}_1^n,\mathrm{div}\boldsymbol{w}^n\right)\right]+\frac{1}{\epsilon}\left(\mathrm{div}_h\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t},\mathrm{div}_h(\boldsymbol{w}^{n-1}-\iota_h\boldsymbol{w}^{n-1})\right)\\ \le&\frac{1}{\epsilon\Delta t}\left[(\mathrm{div}_h\boldsymbol{e}_1^{n-1},\mathrm{div}\boldsymbol{w}^{n-1})-\left(\mathrm{div}_h\boldsymbol{e}_1^n,\mathrm{div}\boldsymbol{w}^n\right)\right]+C_\delta\frac{h^2}{\epsilon}\left\|\mathrm{div}_h\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t}\right\|^2+\delta\frac{1}{\epsilon}\|\mathrm{div}\boldsymbol{w}^{n-1}\|_1^2.\end{aligned}$$

Next the approximation property (2.3) of $\iota_h$, Cauchy-Schwarz and Young's inequality give

$$\mathrm{I}_3^n\le C_\delta h^2\left(\frac{1}{\epsilon}\|\mathrm{div}_h\boldsymbol{e}_1^n\|^2+\|\nabla_h\boldsymbol{e}_1^n\|^2\right)+\delta\left(\frac{1}{\epsilon}\|\mathrm{div}\boldsymbol{w}^{n-1}\|_1^2+\|\boldsymbol{w}^{n-1}\|_2^2\right).$$

It remains to bound $\mathrm{I}_4^n$ and $\mathrm{I}_5^n$. By Lemma 2.1 and Young's inequality, we arrive at

$$\mathrm{I}_4^n\le C_\delta h^2\|\nabla_h\boldsymbol{e}_1^n\|^2+\delta\left(\|\boldsymbol{w}^{n-1}\|_2^2+\frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right\|_1^2+\frac{1}{\epsilon^2}\|\mathrm{div}\boldsymbol{w}^{n-1}\|_1^2\right).$$

Since $\boldsymbol{w}^{n-1}$ is continuous on all interior edges of the grid $\mathscr{T}_h$, we get

$$\begin{aligned}\mathrm{I}_5^n=&X_1\left(\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1}-\boldsymbol{w}^{n-1}\right)+\frac{1}{\epsilon}X_2\left(\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t},\iota_h\boldsymbol{w}^{n-1}-\boldsymbol{w}^{n-1}\right)\\ &+\frac{\alpha}{\epsilon}X_2\left(\mathrm{div}\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1}-\boldsymbol{w}^{n-1}\right)-X_3\left(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\iota_h\boldsymbol{w}^{n-1}-\boldsymbol{w}^{n-1}\right)\\ \le&C_\delta h^4\left(\|\boldsymbol{u}^n\|_2^2+\frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2+\frac{1}{\epsilon^2}\|\mathrm{div}\boldsymbol{u}^n\|_1^2+\|\boldsymbol{u}^{n-1}\|_2^2\|\boldsymbol{u}^n\|_1^2\right)+\delta\|\boldsymbol{w}^{n-1}\|_2^2,\end{aligned}$$

where the last line have used Lemma 2.1, the approximation property (2.3) of $\iota_h$ and Young's inequality.

By direct computation, we have

$$\mathrm{div}_h\boldsymbol{e}_1^0=\mathrm{div}_h(\boldsymbol{u}_0-\iota_h\boldsymbol{u}_0)=(\mathbb{I}-P_h)\mathrm{div}\boldsymbol{u}_0=0,$$

see [12, Lemma 3.1], where $P_h$ is piecewise-constant $L^2(\Omega)$ projection. Then summing up (4.9) over $n$ from $N$ to 1, meanwhile noticing $\boldsymbol{w}^N=0$ and the preliminary estimates on $\{\mathrm{I}_i^n\}_{i=1}^5$ above, together with Poincáre inequality, Cauchy Schwarz and Young's inequality, we have

$$\begin{aligned}\sum_{n=N}^{1}\|\boldsymbol{e}_1^n\|^2\Delta t\le&(\boldsymbol{e}_1^0,\boldsymbol{w}^0)+Ch^4\sum_{n=N}^{1}\left\|\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t}\right\|^2\Delta t+C_\delta h^2\left(\frac{1}{\epsilon}\sum_{n=N}^{1}\left\|\mathrm{div}_h\frac{\boldsymbol{e}_1^n-\boldsymbol{e}_1^{n-1}}{\Delta t}\right\|^2\Delta t\right)\\&+C_\delta h^2\sum_{n=N}^{1}\left(\frac{1}{\epsilon}\|\mathrm{div}_h\boldsymbol{e}_1^n\|^2+\|\nabla_h\boldsymbol{e}_1^n\|^2\right)\Delta t+C_\delta h^2\sum_{n=N}^{1}\|\nabla_h\boldsymbol{e}_1^n\|^2\\&+C_\delta h^4\sum_{n=N}^{1}\left(\|\boldsymbol{u}^n\|_2^2+\frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{u}^n-\boldsymbol{u}^{n-1}}{\Delta t}\right\|_1^2+\frac{1}{\epsilon^2}\|\mathrm{div}\boldsymbol{u}^n\|_1^2+\|\boldsymbol{u}^{n-1}\|_2^2\|\boldsymbol{u}^n\|_1^2\right)\\&+\delta\sum_{n=N}^{1}\left(\|\boldsymbol{w}^{n-1}\|_2^2+\frac{1}{\epsilon^2}\left\|\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right\|_1^2\Delta t+\frac{1}{\epsilon^2}\|\mathrm{div}\boldsymbol{w}^{n-1}\|_1^2\right)\Delta t\\\le& C_\delta h^4+\delta\|\nabla\boldsymbol{w}^0\|^2+C\delta\sum_{n=N}^{1}\|\boldsymbol{e}_1^n\|^2\Delta t\le C_\delta h^4+C\delta\sum_{n=N}^{1}\|\boldsymbol{e}_1^n\|^2\Delta t,\end{aligned}$$

where we have used the estimate (A.2), Lemmas 4.1 and 3.3. By letting $C\delta\le\frac{1}{2}$, the desired estimate follows. We complete the proof. □

Now we turn to derive the error bound on the error $\boldsymbol{e}_2^n$.

**Lemma 4.4.** *For any $\boldsymbol{v}_h\in\boldsymbol{V}_h$, there holds*

$$\|\boldsymbol{v}_h\|_{L^\infty}+\|\nabla_h\boldsymbol{v}_h\|_{L^3}\le C\|\nabla_h\boldsymbol{v}_h\|^{\frac{1}{2}}\|\Delta_h\boldsymbol{v}_h\|^{\frac{1}{2}}.$$

*Proof.* See [10, Lemma 4.4]. □

**Lemma 4.5.** *Let $\{\boldsymbol{u}_*^n\}_{n=0}^N$ be the solution of problem* (4.1). *Then for sufficiently small $\epsilon$, we have*

$$\max_{n=0,1,\cdots N}\|\boldsymbol{u}_*^n\|_{L^\infty}+\max_{n=0,1,\cdots N}\|\nabla_h\boldsymbol{u}_*^n\|_{L^3}\le C.$$

*Proof.* We only need to prove that $\|\nabla_h\boldsymbol{u}_*^n\|$ and $\|\Delta_h\boldsymbol{u}_*^n\|$ are uniformly bounded for every $n\in\{0,1,\cdots N\}$ due to Lemma 4.4. From Lemmas 4.1 and 3.3, we know that

$$\|\nabla_h\boldsymbol{u}_*^n\|\le\|\nabla_h\boldsymbol{e}_1^n\|+\|\nabla\boldsymbol{u}^n\|\le C.$$

Next we consider $\|\Delta_h\boldsymbol{u}_*^n\|$. By triangle inequality we have

$$\|\Delta_h\boldsymbol{u}_*^n\|\le\|\Delta_h\boldsymbol{e}_1^n\|+\|\Delta\boldsymbol{u}^n\|.$$

By definition of discrete Laplacian $\Delta_h$, inverse property (2.2) of $\boldsymbol{V}_h$ and Lemma 4.1, we have

$$\|\Delta_h\boldsymbol{e}_1^n\|^2=-(\nabla_h\boldsymbol{e}_1^n,\nabla_h\Delta_h\boldsymbol{e}_1^n)\le\|\nabla_h\boldsymbol{e}_1^n\|\|\nabla_h\Delta_h\boldsymbol{e}_1^n\|\le Ch\times\frac{C}{h}\|\Delta_h\boldsymbol{e}_1^n\|,$$

and hence $\|\Delta_h\boldsymbol{e}_1^n\|\le C$. This and Lemma 3.3 derive the desired estimate. □

**Lemma 4.6.** *The following properties hold*

$$|b_h(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})|\le C\|\boldsymbol{u}\|_2\|\nabla_h\boldsymbol{w}\|(\|\boldsymbol{v}\|+h\|\nabla_h\boldsymbol{v}\|),\qquad \forall\boldsymbol{u}\in\boldsymbol{H}^2(\Omega)\cap\boldsymbol{H}_0^1(\Omega),\quad \boldsymbol{v},\boldsymbol{w}\in\boldsymbol{H}, \tag{i}$$

$$|b_h(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})|\le C\|\boldsymbol{u}\|\|\nabla_h\boldsymbol{w}\|(\|\boldsymbol{v}\|_{L^\infty}+\|\nabla_h\boldsymbol{v}\|_{L^3}),\quad \forall\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\boldsymbol{H}. \tag{ii}$$

*Proof.* See [12, Lemma 4.6] for the assertion (i). Using the definition of trilinear form $b_h(\cdot,\cdot,\cdot)$, meanwhile applying Hölder's inequality and Sobolev inequality could derive the assertion (ii). □

The next lemma provides a main result on the error $\boldsymbol{e}_2^n$.

**Lemma 4.7.** *For sufficiently small $\epsilon$, there holds*

$$\max_{n=1,\cdots,N}\|\boldsymbol{e}_2^n\|+h\max_{n=1,\cdots,N}\|\nabla_h\boldsymbol{e}_2^n\|\le Ch^2.$$

*Proof.* To handle with the nonlinear part, we first split it into

$$\begin{aligned}&b_h(\boldsymbol{u}^{n-1},\boldsymbol{u}^n,\boldsymbol{v}_h)-b_h(\boldsymbol{u}_h^{n-1},\boldsymbol{u}_h^n,\boldsymbol{v}_h)\\ =&b_h(\boldsymbol{u}^{n-1},\boldsymbol{e}_1^n,\boldsymbol{v}_h)+b_h(\boldsymbol{u}^{n-1},\boldsymbol{u}_*^n,\boldsymbol{v}_h)+b_h(\boldsymbol{u}_h^{n-1},\boldsymbol{e}_2^n,\boldsymbol{v}_h)-b_h(\boldsymbol{u}_h^{n-1},\boldsymbol{u}_*^n,\boldsymbol{v}_h)\\ =&b_h(\boldsymbol{u}^{n-1},\boldsymbol{e}_1^n,\boldsymbol{v}_h)+b_h(\boldsymbol{u}_h^{n-1},\boldsymbol{e}_2^n,\boldsymbol{v}_h)+b_h(\boldsymbol{e}_1^{n-1},\boldsymbol{u}_*^n,\boldsymbol{v}_h)+b_h(\boldsymbol{e}_2^{n-1},\boldsymbol{u}_*^n,\boldsymbol{v}_h).\end{aligned}$$

Let $\boldsymbol{v}_h=\boldsymbol{e}_2^n$ in the weak formulation (4.3). Since $b_h(\boldsymbol{u}_h^{n-1},\boldsymbol{e}_2^n,\boldsymbol{e}_2^n)=0$, then we have

$$\begin{aligned}&\frac{1}{\Delta t}\big(\|\boldsymbol{e}_2^n\|^2-\|\boldsymbol{e}_2^{n-1}\|^2\big)+\frac{1}{2\epsilon\Delta t}\big(\|\mathrm{div}_h\boldsymbol{e}_2^n\|^2-\|\mathrm{div}_h\boldsymbol{e}_2^{n-1}\|^2\big)\\ &\quad+\frac{\alpha}{\epsilon}\|\mathrm{div}_h\boldsymbol{e}_2^n\|^2+\|\nabla_h\boldsymbol{e}_2^n\|^2\\ \le&-b_h(\boldsymbol{u}^{n-1},\boldsymbol{e}_1^n,\boldsymbol{e}_2^n)-b_h(\boldsymbol{e}_1^{n-1},\boldsymbol{u}_*^n,\boldsymbol{e}_2^n)-b_h(\boldsymbol{e}_2^{n-1},\boldsymbol{u}_*^n,\boldsymbol{e}_2^n)\\ =:&\mathrm{I}_1^n+\mathrm{I}_2^n+\mathrm{I}_3^n.\end{aligned} \tag{4.10}$$

It suffices to bound $\{\mathrm{I}_i^n\}_{i=1}^3$ separately. From Lemmas 4.6 and 3.3, we get

$$\begin{aligned}\mathrm{I}_1^n&\le C\|\boldsymbol{u}^{n-1}\|_2\|\nabla_h\boldsymbol{e}_2^n\|\big(\|\boldsymbol{e}_1^n\|+h\|\nabla_h\boldsymbol{e}_1^n\|\big)\\ &\le C\|\boldsymbol{e}_1^n\|^2+Ch^2\|\nabla_h\boldsymbol{e}_1^n\|^2+\frac14\|\nabla_h\boldsymbol{e}_2^n\|^2,\end{aligned}$$

where we have used Young's inequality. Moreover, there also holds (due to Lemma 4.5)

$$\mathrm{I}_2^n\le C\|\boldsymbol{e}_1^{n-1}\|\|\nabla_h\boldsymbol{e}_2^n\|(\|\boldsymbol{u}_*^n\|_{L^\infty}+\|\nabla_h\boldsymbol{u}_*^n\|_{L^3})\le C\|\boldsymbol{e}_1^{n-1}\|^2+\frac14\|\nabla_h\boldsymbol{e}_2^n\|^2,$$

$$\mathrm{I}_3^n\le C\|\boldsymbol{e}_2^{n-1}\|\|\nabla_h\boldsymbol{e}_2^n\|(\|\boldsymbol{u}_*^n\|_{L^\infty}+\|\nabla_h\boldsymbol{u}_*^n\|_{L^3})\le C\|\boldsymbol{e}_2^{n-1}\|^2+\frac14\|\nabla_h\boldsymbol{e}_2^n\|^2.$$

Summing up (4.10) over $n$ from 1 to $k\in\{1,2\cdots N\}$ and noting $\boldsymbol{e}_2^0=\boldsymbol{0}$ yield

$$\begin{aligned}&\|\boldsymbol{e}_2^k\|^2+\frac{1}{\epsilon}\|\mathrm{div}_h\boldsymbol{e}_2^k\|^2+\frac{1}{\epsilon}\sum_{n=0}^{k}\|\mathrm{div}_h\boldsymbol{e}_2^n\|^2\Delta t+\sum_{n=1}^{k}\|\nabla_h\boldsymbol{e}_2^n\|^2\Delta t\\ \le&C\sum_{n=1}^{k}\|\boldsymbol{e}_1^n\|^2\Delta t+Ch^2\sum_{n=0}^{k}\|\nabla_h\boldsymbol{e}_1^n\|^2\Delta t+C\sum_{n=0}^{k}\|\boldsymbol{e}_2^n\|^2\Delta t\\ \le&Ch^4+C\sum_{n=0}^{k}\|\boldsymbol{e}_2^n\|^2\Delta t,\end{aligned}$$

where the last step uses Lemmas 4.1 and 4.3. By discrete Gronwall's inequality, we obtain

$$\max_{n=1,2,\cdots,N}\|\boldsymbol{e}_2^n\|\le Ch^2.$$

Since $\boldsymbol{e}_2^n\in\boldsymbol{V}_h$ for $n=1,2,\cdots,N$, the inverse property (2.2) implies

$$\|\nabla_h\boldsymbol{e}_2^n\|\le\frac{C}{h}\|\boldsymbol{e}_2^n\|\le Ch.$$

The proof of the lemma is completed. □

Now from Lemmas 4.1, 4.3 and 4.7, we have the following result immediately.

**Theorem 4.1.** *Let* $\{\boldsymbol{u}^n\}_{n=0}^N$ *be the solution of time discretization scheme* (1.4a)-(1.4b) *and* $\{\boldsymbol{u}_h^n\}_{n=0}^N$ *the solution of full discretization scheme* (1.5a)-(1.5b). *Assume that* (A1)-(A4) *holds. For sufficiently small* $\epsilon$, *we have*

$$\max_{n=1,2,\cdots,N}h^2\|\nabla_h(\boldsymbol{u}^n-\boldsymbol{u}_h^n)\|^2+\sum_{n=1}^{N}\|\boldsymbol{u}^n-\boldsymbol{u}_h^n\|^2\Delta t\le Ch^4.$$

In [13], time discretization scheme (1.4a)-(1.4b) is studied and an error estimate between the scheme and Navier-Stokes equations (1.1a)-(1.1c) is given. Combining them together, we have the error estimation to the full discretization scheme (1.5a)-(1.5b).

**Theorem 4.2.** *Let* $(\boldsymbol{u},p)$ *be the solution of unsteady Navier-Stokes equations* (1.1a)-(1.1c) *and* $\{\boldsymbol{u}_{sh}^n\}_{n=0}^N$ *the solution of full discretization scheme* (1.5a)-(1.5b). *Assume that* (A1)-(A4) *holds. We further assume that*

$$\int_0^T\|p_0\|_1^2dt\le2+2\int_0^T\|p\|_1^2dt,$$

*and $\epsilon$ is small enough. Then for $s=1,2,\cdots$, we have*

$$\max_{n=1,\cdots,N}\|\nabla_h(\boldsymbol{u}_{sh}^n-\boldsymbol{u}(\cdot,n\Delta t))\|^2\le C(h^2+\Delta t^2+\epsilon^{2s}),$$
$$\sum_{n=1}^N\|\boldsymbol{u}_{sh}^n-\boldsymbol{u}(\cdot,n\Delta t)\|^2\Delta t\le C(h^4+\Delta t^2+\epsilon^{2s}).$$

**Remark 4.1.** The result provides an optimal error bounds in both $\boldsymbol{H}^1$ and $\boldsymbol{L}^2$, with respect to the chosen small penalty $\epsilon$ and discretization parameters. It has essentially improved the estimation using the conforming $P_1$ element approximation, which depends inversely on the penalty $\epsilon$.

# 5 Numerical experiments

In this section we present numerical experiments of the sequential regularization method for two dimensional Navier-Stokes equations using nonconforming $P_1$ finite element discretization. We choose $\alpha_1=1$, $\alpha_2=1$, $\nu=1$ in all examples.

**Example 5.1.** Let $\Omega=(0,1)^2$, $T=1$, $\epsilon=$1.00e-2, $s=5$ and initial guess $p_0(\boldsymbol{x},t)\equiv 0$ in $\Omega\times(0,T)$. The forcing term $\boldsymbol{f}_{ext}$ is chosen such that the exact velocity and pressure are given by, respectively,

$$u_1(\boldsymbol{x},t)=\frac{1}{2}x^2(1-x)^2(4y^3-6y^2+2y)e^t,$$
$$u_2(\boldsymbol{x},t)=-\frac{1}{2}y^2(1-y)^2(4x^3-6x^2+2x)e^t,$$
$$p(\boldsymbol{x},t)=(2x-1)(2y-1)e^t.$$

To investigate the convergence behavior of the discrete approximation $\boldsymbol{u}_{sh}^n$ in $H^1$ and $L^2$ norms, we define two different metrics as follows:

$$e_{sH}:=\sum_{n=1}^N\|\nabla_h(\boldsymbol{u}_{sh}^n-\boldsymbol{u}(\cdot,n\Delta t))\|^2\Delta t \quad \text{and} \quad e_{sL}:=\sum_{n=1}^N\|\boldsymbol{u}_{sh}^n-\boldsymbol{u}(\cdot,n\Delta t)\|^2\Delta t,$$

respectively. The corresponding grids, velocity fields, and numerical results are presented in Fig. 1 and Table 1. Steady convergence of both $e_{sH}$ and $e_{sL}$ can be observed from the last column of Table 1, and their convergence rates confirm to theoretical results.

**Example 5.2.** Let $\Omega=(0,\pi)^2$, $T=1$, $\epsilon=$1.00e-2, $s=5$. Set initial guess

$$p_0(\boldsymbol{x},t)\equiv 0 \quad \text{and} \quad g=\frac{1}{8}\sin^2(x)\sin^2(y)e^t$$

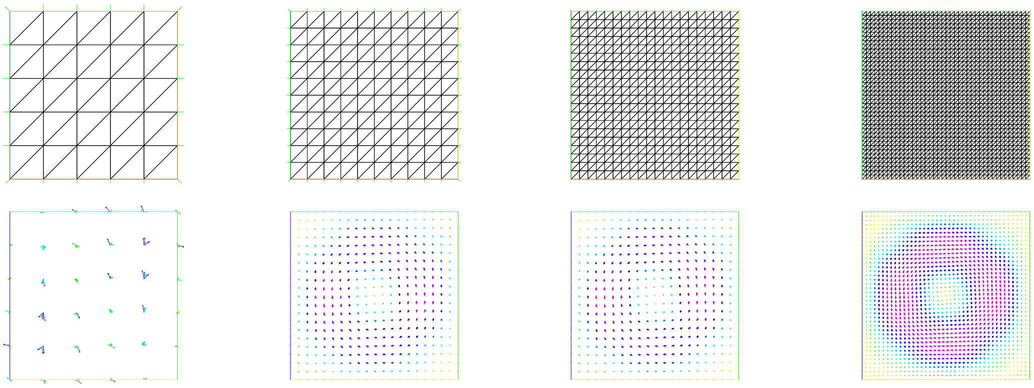

Figure 1: A sequence of grids and their corresponding fields of the velocity at time $T=1$ for Example 5.1.

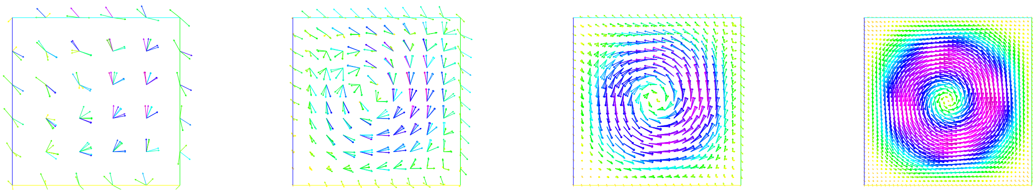

Figure 2: The fields of the velocity at time $T=1$ for Example 5.2.

in $\Omega\times(0,T)$. The forcing term $\boldsymbol{f}_{ext}$ is chosen such that the exact velocity $\boldsymbol{u}(\boldsymbol{x},t)=$ curl $g$ and the pressure $p(\boldsymbol{x},t)=xye^t$.

The computational grids are generated by the same strategy as in the aforementioned Example 5.1. The velocity fields and numerical results for Example 5.2 are presented in Fig. 2 and Table 2. Clearly, the convergence behavior of both cases confirms the main results.

Table 1: Numerical results for Example 5.1.

| # Sdofs | 170 | 640 | 2480 | 9760 | rate |
|---|---|---|---|---|---|
| $e_{sH}$ | 1.46e-1 | 8.30e-2 | 4.29e-2 | 2.14e-2 | 1.00 |
| $e_{sL}$ | 1.18e-2 | 3.79e-3 | 1.06e-3 | 2.76e-4 | 1.94 |

Table 2: Numerical results for Example 5.2.

| # Sdofs | 170 | 640 | 2480 | 9760 | rate |
|---|---|---|---|---|---|
| $e_{sH}$ | 3.14e0 | 1.66e0 | 8.29e-1 | 4.10e-1 | 0.98 |
| $e_{sL}$ | 8.96e-1 | 2.43e-1 | 6.35e-2 | 1.62e-2 | 1.97 |

# Appendix A: A crucial priori regularity on duality problem

In this appendix, we collect the proof of Lemma 4.2.

*Proof.* Define by

$$\begin{aligned} A\boldsymbol{w}^{n-1} &= -\frac{1}{\epsilon}\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t} - \frac{\alpha}{\epsilon}\mathrm{div}\boldsymbol{w}^{n-1} - \Delta\boldsymbol{w}^{n-1}, \\ p^{n-1} &= -\frac{1}{\epsilon}\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t} - \frac{\alpha}{\epsilon}\mathrm{div}\boldsymbol{w}^{n-1}. \end{aligned}$$

Let $g^{n-1} = \mathrm{div}\boldsymbol{w}^{n-1}$. Then the second identity could be rewritten as

$$g^{n-1} = \frac{1}{(1+\alpha\Delta t)}g^n - \frac{\epsilon\Delta t}{(1+\alpha\Delta t)}p^{n-1},$$

which implies that

$$\|g^{n-1}\|_1 \le \|g^n\|_1 + \epsilon\Delta t\|p^{n-1}\|_1.$$

Applying mathematics induction for the inequality and noting

$$g^N = \mathrm{div}\boldsymbol{w}^N = 0,$$

then there holds for any $k \in \{N, N-1, \cdots, 1\}$

$$\|g^{k-1}\|_1^2 \le \epsilon^2\Delta t^2\left(\sum_{n=k}^{N}\|p^{n-1}\|_1\right)^2 \le \epsilon^2 T\sum_{n=N}^{k}\|p^{n-1}\|_1^2\Delta t, \tag{A.1}$$

where the last step uses Cauchy-Schwarz inequality.

Now upon the standard result of Stokes equations (cf. [16]), we have the a priori regularity

$$\|\boldsymbol{w}^{n-1}\|_2^2 + \|p^{n-1}\|_1^2 \le C\big(\|A\boldsymbol{w}^{n-1}\|^2 + \|g^{n-1}\|_1^2\big).$$

Multiplying $\Delta t$ on both sides and summing up the estimate over $n$ from $N$ to 1 above, together with the elementary estimate (A.1), lead to

$$\begin{aligned}
&\sum_{n=N}^{1}\|\boldsymbol{w}^{n-1}\|_2^2\Delta t+\sum_{n=N}^{1}\|p^{n-1}\|_1^2\Delta t\\
\le& C\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t+C\sum_{n=N}^{1}\|g^{n-1}\|_1^2\Delta t\\
\le& C\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t+C\epsilon^2T\sum_{n=N}^{1}\sum_{k=N}^{n}\|p^{k-1}\|_1^2\Delta t^2\\
\le& C\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t+C\epsilon^2T^2\sum_{n=N}^{1}\|p^{n-1}\|_1^2\Delta t.
\end{aligned}$$

Let $\epsilon$ be small enough such that $C\epsilon^2T^2\le\frac{1}{2}$. Then we have

$$\sum_{n=N}^{1}\|\boldsymbol{w}^{n-1}\|_2^2\Delta t+\sum_{n=N}^{1}\|p^{n-1}\|_1^2\Delta t\le C\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t.$$

Consequently, there holds for the function $\mathrm{div}\boldsymbol{w}^{n-1}$

$$\frac{1}{\epsilon^2}\sum_{n=N}^{1}\|\mathrm{div}\boldsymbol{w}^{n-1}\|_1^2\Delta t\le T\sum_{k=N}^{n}\|p^{k-1}\|_1^2\Delta t\le C\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t,$$

and further the definition of $p^{n-1}$ and triangle inequality imply

$$\begin{aligned}
&\frac{1}{\epsilon^2}\sum_{n=N}^{1}\left\|\mathrm{div}\frac{\boldsymbol{w}^n-\boldsymbol{w}^{n-1}}{\Delta t}\right\|_1^2\Delta t=\sum_{n=N}^{1}\left\|p^{n-1}+\frac{\alpha}{\epsilon}\mathrm{div}\boldsymbol{w}^{n-1}\right\|_1^2\Delta t\\
\le& C\sum_{n=N}^{1}\|p^{n-1}\|_1^2\Delta t+C\frac{\alpha^2}{\epsilon^2}\sum_{n=N}^{1}\|\mathrm{div}\boldsymbol{w}^{n-1}\|_1^2\Delta t\le C\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t.
\end{aligned}$$

It suffices to prove

$$\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t\le C\sum_{n=N}^{1}\|\boldsymbol{e}_1^n\|^2\Delta t.$$

Multiplying $\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}$ and integrating over the domain $\Omega$ on both sides of (4.5) derive

$$\begin{aligned}
&\left\|\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right\|^2+\frac{1}{\epsilon}\left\|\mathrm{div}\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right\|^2+\frac{\alpha}{2\epsilon\Delta t}\left(\|\mathrm{div}\boldsymbol{w}^{n-1}\|^2-\|\mathrm{div}\boldsymbol{w}^n\|^2\right)\\
&+\frac{1}{2\Delta t}\left(\|\nabla\boldsymbol{w}^{n-1}\|^2-\|\nabla\boldsymbol{w}^n\|^2\right)\le\left(\boldsymbol{e}_1^n,\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right).
\end{aligned}$$

Summing up from $n=N$ to 1 and noticing that $\boldsymbol{w}^N=0$, meanwhile using Cauchy-Schwarz and Young's inequality, we could arrive at

$$\max_{n=1,2,\cdots,N}\|\nabla\boldsymbol{w}^{n-1}\|^2+\sum_{n=N}^{k}\left\|\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right\|^2\Delta t\leq C\sum_{n=N}^{1}\|\boldsymbol{e}_1^n\|^2\Delta t. \tag{A.2}$$

Next we input the test function $A\boldsymbol{w}^{n-1}$ into (4.5) and get

$$\begin{aligned}\|A\boldsymbol{w}^{n-1}\|^2&=(\boldsymbol{e}_1^n,A\boldsymbol{w}^{n-1})-\left(\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t},A\boldsymbol{w}^{n-1}\right)\\&\leq\|\boldsymbol{e}_1^n\|^2+\left\|\frac{\boldsymbol{w}^{n-1}-\boldsymbol{w}^n}{\Delta t}\right\|^2+\frac{1}{2}\|A\boldsymbol{w}^{n-1}\|^2,\end{aligned}$$

and hence we obtain

$$\sum_{n=N}^{1}\|A\boldsymbol{w}^{n-1}\|^2\Delta t\leq C\sum_{n=N}^{1}\|\boldsymbol{e}_1^n\|^2\Delta t.$$

Combining with all estimates above, the desire estimates follows. We complete the proof. □

# Acknowledgements

We would like the thank the anonymous referees and associated editor for their useful comments and suggestions, which have led to considerable improvements in the paper. This work is supported by the National Key Research and Development Program of China (No. 2020YFA0714200), by the National Science Foundation of China (No. 12371424). The numerical calculations have been done at the Supercomputing Center of Wuhan University.

# References

[1] Uri Ascher and Ping Lin, Sequential regularization methods for higher index daes with constraint singularities: The linear index-2 case, SIAM J. Numer. Anal., 33(5) (1996), 1921–1940.
[2] Uri Ascher and Ping Lin, Sequential regularization methods for nonlinear higher-index daes, SIAM J. Sci. Comput., 18(1) (1997), 160–181.
[3] Joachim Baumgarte, Stabilization of constraints and integrals of motion in dynamical systems, Comput. Methods Appl. Mech. Eng., 1(1) (1972), 1–16.

[4] B. Brefort, J. Mo Ghidaglia, and R. Temam, Attractors for the penalized Navier–Stokes equations, SIAM J. Math. Anal., 19(1) (1988), 1–21.

[5] Kathryn Eleda Brenan, Stephen L. Campbell, and Linda Ruth Petzold, Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations, SIAM, 1995.

[6] Susanne C. Brenner and Li-Yeng Sung, Linear finite element methods for planar linear elasticity, Math. Comput., 59(200) (1992), 321–338.

[7] Franco Brezzi and Michel Fortin, Mixed and Hybrid Finite Element Methods, volume 15, Springer Science & Business Media, 2012.

[8] Alexandre Joel Chorin, Numerical solution of the Navier-Stokes equations, Math. Comput., 22(104) (1968), 745–762.

[9] Michel Crouzeix and P.-A. Raviart, Conforming and nonconforming finite element methods for solving the stationary stokes equations I, ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, 7(R3) (1973), 33–75.

[10] John G. Heywood and Rolf Rannacher, 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(2) (1982), 275–311.

[11] Ping Lin, A sequential regularization method for time-dependent incompressible Navier–Stokes equations, SIAM J. Numer. Anal., 34(3) (1997), 1051–1071.

[12] Xiliang Lu and Ping Lin, Error estimate of the $p_1$ nonconforming finite element method for the penalized unsteady Navier-Stokes equations, Numerische Mathematik, 115(2) (2010), 261–287.

[13] Xiliang Lu, Ping Lin, and Jianguo Liu, Analysis of a sequential regularization method for the unsteady Navier-Stokes equations, Math. Comput., 77(263) (2008), 1467–1494.

[14] Jie Shen, On error estimates of the penalty method for unsteady Navier–Stokes equations, SIAM J. Numer. Anal., 32(2) (1995), 386–403.

[15] Roger Temam, Sur l'approximation de la solution des équations de navier-stokes par la méthode des pas fractionnaires (ii), Arch. Rational Mech. Anal., 33(5) (1969), 377–385.

[16] Roger Temam, Navier-Stokes Equations: Theory and Numerical Analysis, volume 343, American Mathematical Soc., 2001.

[17] Xiliang Lu, Error Analysis for Incompressible Viscous Flow Based on a Sequential Regularization Formulation, PhD thesis, Ph. D. thesis, Department of Mathematics, National University of Singapore, 2006.