

Iterative penalty finite element methods for the steady incompressible magnetohydrodynamic problem

Jien Deng¹ · Zhenzhen Tao¹ · Tong Zhang^{1,2}

Received: 20 June 2015 / Revised: 31 January 2016 / Accepted: 2 February 2016 /
Published online: 23 February 2016
© SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional 2016

Abstract In the paper we consider the one-level and two-level iterative penalty finite element methods for the steady incompressible magnetohydrodynamic problem based on the iteration of pressure with a factor of penalty parameter. Firstly, the \mathbf{H}^1 and L^2 error estimates of numerical solutions of one-level iterative penalty finite element method are provided. Secondly, the stability and convergence of two-level iterative penalty finite element method are analyzed. Finally, some numerical results are provided to verify the effectiveness of the developed numerical schemes.

Keywords Iterative penalty method · Two-level method · Steady incompressible MHD problem · Stability · Convergence

Mathematics Subject Classification 35Q30 · 65M60 · 65N30

1 Introduction

In this paper, let $\Omega \subset R^d$ ($d = 2$ or 3) be a convex polygonal/polyhedral domain (see [Gunzburger et al. 1991, 2004](#)). We consider the following steady incompressible MHD problem:

Communicated by Jorge X. Velasco.

This work was supported by CAPES and CNPq, Brazil (No. 88881.068004/2014.01), the NSF of China (No. 11301157), the Doctor Fund of Henan Polytechnic University (B2012-098) and the Foundation of Distinguished Young Scientists of Henan Polytechnic University (J2015-05).

✉ Jien Deng
dengjien@hpu.edu.cn

Tong Zhang
zhangtong0616@163.com

¹ School of Mathematics and Information Science, Henan Polytechnic University, Jiaozuo 454003, People's Republic of China

² Departamento de Matemática, Universidade Federal do Paraná, Curitiba, PR 81531-990, Brazil

$$\begin{cases} -R_e^{-1} \Delta \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - S_c \operatorname{curl} \mathbf{B} \times \mathbf{B} = \mathbf{f} & \text{in } \Omega, \\ S_c R_m^{-1} \operatorname{curl}(\operatorname{curl} \mathbf{B}) - S_c \operatorname{curl}(\mathbf{u} \times \mathbf{B}) = \mathbf{g} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \nabla \cdot \mathbf{B} = 0 & \text{in } \Omega, \end{cases} \quad (1.1)$$

subject to the boundary conditions

$$\begin{cases} \mathbf{u} = 0 & \text{on } \partial\Omega, \\ \mathbf{B} \cdot \mathbf{n} = 0 & \text{on } \partial\Omega, \\ \mathbf{n} \times \operatorname{curl} \mathbf{B} = 0 & \text{on } \partial\Omega, \end{cases} \quad (1.2)$$

where \mathbf{u} is the velocity field, \mathbf{B} denotes the magnetic field, \mathbf{f} and \mathbf{g} are the source terms. \mathbf{n} is outward normal unit vector of $\partial\Omega$, p is the hydrodynamic pressure, R_e , R_m and S_c are the hydrodynamic Reynolds number, magnetic Reynolds number and coupling number, respectively.

The steady incompressible MHD problem can be used to describe the interaction between a viscous, incompressible, electrically conducting field and an external magnetic field. Namely, the steady incompressible MHD problem is a coupled system, which is composed of Navier–Stokes equations of fluid dynamics and Maxwell’s equations that couple Lorentz’s force with Ohm’s law. We refer to [Hughes and Young \(1966\)](#) and [Moreau et al. \(1990\)](#) for comprehensive accounts of the physical background of MHD problem. Several papers have been devoted to the design and the analysis of numerical schemes for the MHD problem. For example, we can refer to [Gunzburger et al. \(1991, 2004\)](#) for the existence and uniqueness of the solutions, [Discacciati \(2008\)](#) for numerical approximation of the steady MHD problem, [Hasler et al. \(2004\)](#) and [Schözau \(2004\)](#) for the mixed finite element method (FEM), and [Dong et al. \(2014\)](#) and [Tao and Zhang \(2015\)](#) for the iterative method and so on.

The first main difficulty of solving the MHD problem is the nonlinear terms $\mathbf{u} \cdot \nabla \mathbf{u}$, $\operatorname{curl} \mathbf{B} \times \mathbf{B}$ and $\operatorname{curl}(\mathbf{u} \times \mathbf{B})$. Two-level method is an efficient numerical scheme for the nonlinear terms, and this method was pioneered by [Marion and Xu \(1995\)](#) and [Xu \(1996\)](#). The main idea of two level method is to find an initial approximation on a coarse mesh firstly, and then to solve a linear problem by using the coarse mesh solution on a fine mesh. It is a good strategy to decrease the computational cost. Therefore, two-level method has been wildly studied in recent years. For example, we can refer to [Girault and Lions \(2001\)](#), [He \(2003, 2004\)](#) and [Zhang and Yang \(2014\)](#) for the research of the Navier–Stokes equations, the nonlinear parabolic problem ([Zhang 2013](#)) and the natural convection problem ([Zhang et al. 2015a, b](#)). The other main difficulty is that the velocity and the pressure are coupled. Penalty method is a method to overcome this difficulty. Certainly, many researchers have focused on studying penalty method for solving different problems. For example, we can refer to [Dai \(2007\)](#) for the pure Neumann problem, and [An and Shi \(2015\)](#), [Gunzburger \(1989\)](#), [He \(2005\)](#) and [Shen \(1995\)](#) for the incompressible flow. From above mentioned literature, we know that the combination of two-level method and penalty method is quite efficient for solving the nonlinear system. Especially, from the numerical results of [An and Shi \(2015\)](#) and [Qiu et al. \(2014\)](#), we can see that two-level iterative FEM can save much CPU time than one-level iterative FEM with the same convergence order.

In this paper we consider the one-level and two-level iterative penalty FEMs to solve problem (1.1). The penalty parameter ε ($0 < \varepsilon \ll 1$) is set as a real number. For any positive integer k , which is the number of iteration, the error estimates of the one-level iterative penalty FEM solution $((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), p_{\varepsilon\mu}^k)$ are

$$\begin{aligned} \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 + \|p - p_{\varepsilon\mu}^k\|_0 &\leq C(\mu + \varepsilon^{k+1}), \\ \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0 &\leq C(\mu^2 + \mu\varepsilon + \varepsilon^{k+1}), \end{aligned}$$

and the error estimate of two-level iterative penalty FEM solution $((\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h), p_\varepsilon^h)$ is

$$\|(\mathbf{u} - \mathbf{u}_\varepsilon^h, \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 + \|p - p_\varepsilon^h\|_0 \leq C(h + H^2 + \varepsilon H + \varepsilon^{k+1}).$$

Thus, if we choose $\varepsilon = \mathcal{O}(H) = \mathcal{O}(h^{1/2})$, the one-level and two-level iterative penalty FEMs have the same order as the standard Galerkin FEM (see Dong et al. 2014). While from the point of view of numerical tests, we know that the two-level iterative penalty FEM can save a large amount of computational time than one-level iterative penalty FEM with the same order.

The paper is organized as follows: some notations and basic results of problem (1.1) are recalled in Sect. 2; stability and convergence of iterative penalty FEM are presented in Sect. 3; the stability and convergence of two-level iterative penalty FEM are analyzed in Sect. 4; and some numerical experiments are provided to validate the established theoretical analysis in Sect. 5. Finally, some conclusions are given in the last section.

2 Preliminaries

To gain the variational formulation for the steady incompressible MHD flow, we choose the standard Sobolev space $H^j(\Omega) = W^{j,2}(\Omega)$ for any nonnegative integer j with norm $\|v\|_j = (\sum_{|\gamma|=0}^j \|D^\gamma v\|_0^2)^{\frac{1}{2}}$. We use the standard Sobolev space $\mathbf{H}^j(\Omega) = (H^j(\Omega))^d$ with the corresponding norm $\|\mathbf{v}\|_j = (\sum_{i=1}^d \|v_i\|_j^2)^{\frac{1}{2}}$ (see Adams 1975; Girault and Raviart 1986 for more details). Furthermore, we introduce some spaces as follows.

$$\begin{aligned} \mathbf{X} &= \mathbf{H}_0^1(\Omega) = \{\mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v}|_{\partial\Omega} = 0\}, \quad M = L_0^2(\Omega) = \left\{q \in L^2(\Omega) : \int_\Omega q \, dx = 0\right\}, \\ \mathbf{W} &= \{\mathbf{w} \in \mathbf{H}^1(\Omega) : \mathbf{w} \cdot \mathbf{n}|_{\partial\Omega} = 0\}, \quad \mathbf{V} = \{\mathbf{v} \in \mathbf{X} : \nabla \cdot \mathbf{v} = 0 \text{ in } \Omega\}, \\ \mathbf{V}_n &= \{\mathbf{w} \in \mathbf{W} : \nabla \cdot \mathbf{w} = 0 \text{ in } \Omega\}. \end{aligned}$$

With the equivalent norms $\|\nabla \mathbf{w}\|_0$ and $\|\mathbf{w}\|_{\mathbf{H}_0^1(\Omega)}$ of \mathbf{X} , we denote the product space $\mathbf{W}_{0n} = \mathbf{X} \times \mathbf{W}$ equipped with the usual graph norm $\|(\mathbf{w}, \Phi)\|_1, \forall (\mathbf{w}, \Phi) \in \mathbf{W}_{0n}$, where $\|(\mathbf{w}, \Phi)\|_i = (\|\mathbf{w}\|_i^2 + \|\Phi\|_i^2)^{1/2} (i = 0, 1, 2)$. The dual space of $\mathbf{H}_0^1(\Omega)$ is denoted as $\mathbf{H}^{-1}(\Omega)$ which equipped with the norm $\|\cdot\|_{-1}$. In addition, the following two formulas

$$((\mathbf{a} \times \mathbf{b}) \times \mathbf{c}) \cdot \mathbf{d} = (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = -(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{d} \times \mathbf{c}),$$

and

$$\int_\Omega (\nabla \times \Phi) \cdot \Psi \, dx = - \int_{\partial\Omega} (\Phi \times \mathbf{n}) \cdot \Psi \, ds + \int_\Omega \Phi \cdot (\nabla \times \Psi) \, dx,$$

imply that

$$\begin{aligned} (\text{curl}(\mathbf{w} \times \Phi), \Psi)_\Omega &= -\langle (\mathbf{w} \times \Phi) \times \mathbf{n}, \Psi \rangle_{\partial\Omega} + (\mathbf{w} \times \Phi, \text{curl}\Psi)_\Omega \\ &= (\mathbf{w} \times \Phi, \text{curl}\Psi)_\Omega = -(\text{curl}\Psi \times \Phi, \mathbf{w})_\Omega, \quad \forall \mathbf{w} \in \mathbf{X}, \Phi, \Psi \in \mathbf{W}, \end{aligned}$$

where $(\cdot, \cdot)_\Omega$ stands for L^2 inner product on the domain Ω . Define the trilinear term as follows:

$$\begin{aligned} a_1(\mathbf{u}, \mathbf{w}, \mathbf{v}) &= \left(\mathbf{u} \cdot \nabla \mathbf{w} + \frac{1}{2}(\nabla \cdot \mathbf{u})\mathbf{w}, \mathbf{v} \right)_\Omega \\ &= \frac{1}{2}(\mathbf{u} \cdot \nabla \mathbf{w}, \mathbf{v})_\Omega - \frac{1}{2}(\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{w})_\Omega, \quad \forall \mathbf{u}, \mathbf{w}, \mathbf{v} \in \mathbf{X}. \end{aligned} \tag{2.1}$$

With above notations, for $\mathbf{f} \in \mathbf{H}^1(\Omega)$, $\mathbf{g} \in L^2(\Omega)^d$, the weak variational formulation of the steady incompressible MHD problem (1.1) reads as: Find $((\mathbf{u}, \mathbf{B}), p) \in \mathbf{W}_{0n} \times M$ such that

$$\begin{aligned} A_0((\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}, \mathbf{B}), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p) + d_0((\mathbf{u}, \mathbf{B}), q) \\ = \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle, \quad \forall ((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n} \times M, \end{aligned} \tag{2.2}$$

where

$$\begin{aligned} A_0((\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) &= R_e^{-1}a_0(\mathbf{u}, \mathbf{v}) + S_c R_m^{-1}b_0(\mathbf{B}, \Psi), \\ A_1((\mathbf{w}, \Phi), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) &= a_1(\mathbf{w}, \mathbf{u}, \mathbf{v}) - c(\mathbf{B}, \Phi, \mathbf{v}) + c(\Psi, \Phi, \mathbf{u}), \\ d_0((\mathbf{v}, \Psi), q) &= (\nabla \cdot \mathbf{v}, q)_\Omega, \quad a_0(\mathbf{u}, \mathbf{v}) = (\nabla \mathbf{u}, \nabla \mathbf{v})_\Omega, \\ b_0(\mathbf{B}, \Psi) &= (\nabla \times \mathbf{B}, \nabla \times \Psi)_\Omega + (\nabla \cdot \mathbf{B}, \nabla \cdot \Psi)_\Omega, \\ c(\mathbf{B}, \Phi, \mathbf{v}) &= S_c(\text{curl} \mathbf{B} \times \Phi, \mathbf{v})_\Omega, \quad \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle = \langle \mathbf{f}, \mathbf{v} \rangle_\Omega + \langle \mathbf{g}, \Psi \rangle_\Omega. \end{aligned}$$

Furthermore, we define

$$\|\mathbf{F}\|_{-1} = \sup_{(\mathbf{0}, \mathbf{0}) \neq (\mathbf{v}, \Psi) \in \mathbf{W}_{0n}} \frac{\langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle}{\|(\mathbf{v}, \Psi)\|_1}.$$

The following properties of trilinear form $a_1(\cdot, \cdot, \cdot)$ are useful to obtain the existence and uniqueness of a solution to problem (2.2) and gain the corresponding convergence (Adams 1975; Girault and Raviart 1986):

$$a_1(\mathbf{u}, \mathbf{v}, \mathbf{w}) = -a_1(\mathbf{u}, \mathbf{w}, \mathbf{v}), \quad \forall \mathbf{w}, \mathbf{u}, \mathbf{v} \in X, \tag{2.3}$$

$$|a_1(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq C_0^2 \|\nabla \mathbf{u}\|_0 \|\nabla \mathbf{v}\|_0 \|\nabla \mathbf{w}\|_0, \quad \forall \mathbf{w}, \mathbf{u}, \mathbf{v} \in \mathbf{X}, \tag{2.4}$$

$$\begin{aligned} |a_1(\mathbf{u}, \mathbf{v}, \mathbf{w})| &\leq \frac{N}{2} \|\mathbf{u}\|_0 (\|\nabla \mathbf{v}\|_0 \|\mathbf{w}\|_{\mathbf{L}^\infty} + \|\mathbf{v}\|_{\mathbf{L}^6} \|\nabla \mathbf{w}\|_{\mathbf{L}^3}), \\ \forall \mathbf{u} &\in \mathbf{L}^2(\Omega), \mathbf{v} \in \mathbf{X}, \mathbf{w} \in \mathbf{L}^\infty(\Omega) \cap \mathbf{X}, \end{aligned} \tag{2.5}$$

$$\begin{aligned} |a_1(\mathbf{u}, \mathbf{v}, \mathbf{w})| &\leq \frac{N}{2} (\|\mathbf{u}\|_{\mathbf{L}^\infty} \|\nabla \mathbf{v}\|_0 + \|\nabla \mathbf{u}\|_{\mathbf{L}^3} \|\mathbf{v}\|_{\mathbf{L}^6}) \|\mathbf{w}\|_0, \\ \forall \mathbf{u} &\in \mathbf{L}^\infty(\Omega) \cap \mathbf{X}, \mathbf{v} \in \mathbf{X}, \mathbf{w} \in \mathbf{L}^2(\Omega), \end{aligned} \tag{2.6}$$

$$\|\mathbf{v}\|_0 \leq \gamma_0 \|\nabla \mathbf{v}\|_0, \quad \|\mathbf{v}\|_{\mathbf{L}^3} \leq C \|\mathbf{v}\|_0^{\frac{1}{2}} \|\nabla \mathbf{v}\|_0^{\frac{1}{2}}, \quad \|\mathbf{v}\|_{\mathbf{L}^6} \leq C \|\nabla \mathbf{v}\|_0, \quad \forall \mathbf{v} \in \mathbf{X}, \tag{2.7}$$

$$\|\mathbf{v}\|_{\mathbf{L}^\infty} \leq C \|\mathbf{v}\|_1^{\frac{1}{2}} \|\mathbf{v}\|_2^{\frac{1}{2}}, \quad \forall \mathbf{v} \in \mathbf{H}^2(\Omega), \tag{2.8}$$

where $N > 0$ is a constant, γ_0 (only dependent on Ω) is a positive constant and C_0 (only dependent on Ω) is an embedding constant of $\mathbf{H}^1(\Omega) \hookrightarrow \mathbf{L}^4(\Omega)$ (see Adams 1975) (\hookrightarrow denotes the continuous embedding), namely

$$\|\mathbf{w}\|_{\mathbf{L}^4} \leq C_0 \|\nabla \mathbf{w}\|_0, \quad \forall \mathbf{w} \in \mathbf{X}.$$

The trilinear form $A_1(\cdot, \cdot, \cdot)$ is skew symmetric with respect to the later two variables, and it satisfies

$$A_1((\mathbf{w}, \Phi), (\mathbf{u}, \mathbf{B}), (\mathbf{u}, \mathbf{B})) = 0, \quad \forall (\mathbf{w}, \Phi), (\mathbf{u}, \mathbf{B}) \in \mathbf{W}_{0n}. \tag{2.9}$$

To obtain the well-posedness of the problem (2.2), we list the coercivity and continuity of $A_0(\cdot, \cdot)$ and the continuity of $A_1(\cdot, \cdot, \cdot)$ (see Gunzburger et al. 1991): for all $(\mathbf{w}, \Phi), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi) \in \mathbf{W}_{0n}$ such that

$$A_0((\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) \leq \max\{R_e^{-1}, (2 + d)S_c R_m^{-1}\} \|(\mathbf{u}, \mathbf{B})\|_1 \|(\mathbf{v}, \Psi)\|_1, \tag{2.10}$$

$$A_0((\mathbf{u}, \mathbf{B}), (\mathbf{u}, \mathbf{B})) \geq \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}, \mathbf{B})\|_1^2, \tag{2.11}$$

$$A_1((\mathbf{w}, \Phi), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) \leq \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{w}, \Phi)\|_1 \|(\mathbf{u}, \mathbf{B})\|_1 \|(\mathbf{v}, \Psi)\|_1, \tag{2.12}$$

where C_1 (only dependent on Ω) is the constant from the following inequality:

$$\|\nabla \times \Psi\|_0^2 + \|\nabla \cdot \Psi\|_0^2 \geq C_1 \|\Psi\|_1^2, \quad \forall \Psi \in \mathbf{W},$$

$\sqrt{2}$ and d come from two inequalities as follows:

$$\|\text{curl} \mathbf{v}\|_0 \leq \sqrt{2} \|\nabla \mathbf{v}\|_0, \quad \|\nabla \cdot \mathbf{v}\|_0 \leq \sqrt{d} \|\nabla \mathbf{v}\|_0.$$

where d is the dimension of the considered domain Ω .

Thanks to (2.3)–(2.8), the following properties of $A_1(\cdot, \cdot, \cdot)$ hold (see Lemma 1 of Dong et al. 2014):

$$\begin{aligned} |A_1((\mathbf{w}, \Phi), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi))| &\leq C\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{w}, \Phi)\|_0 \|(\mathbf{u}, \mathbf{B})\|_2 \|(\mathbf{v}, \Psi)\|_1, \\ \forall (\mathbf{w}, \Phi) \in \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Omega), \quad (\mathbf{u}, \mathbf{B}) \in \mathbf{H}^2(\Omega) \times \mathbf{H}^2(\Omega), \quad (\mathbf{v}, \Psi) \in \mathbf{W}_{0n}, \end{aligned} \tag{2.13}$$

$$\begin{aligned} |A_1((\mathbf{w}, \Phi), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi))| &\leq C\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{w}, \Phi)\|_2 \|(\mathbf{u}, \mathbf{B})\|_1 \|(\mathbf{v}, \Psi)\|_0, \\ \forall (\mathbf{w}, \Phi) \in \mathbf{H}^2(\Omega) \times \mathbf{H}^2(\Omega), \quad (\mathbf{u}, \mathbf{B}) \in \mathbf{W}_{0n}, \quad (\mathbf{v}, \Psi) \in \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Omega). \end{aligned} \tag{2.14}$$

Throughout this paper, the letter $C > 0$ denotes different constant at different places, and C is independent of the mesh size μ and penalty parameter ε .

The bilinear form $d_0(\cdot, \cdot)$ is continuous on $\mathbf{W}_{0n} \times M$, and it satisfies (see Gunzburger et al. 1991):

$$\sup_{(\mathbf{v}, \Psi) \in \mathbf{W}_{0n}} \frac{|d_0((\mathbf{v}, \Psi), q)|}{\|(\mathbf{v}, \Psi)\|_1} \geq \beta_0 \|q\|_0, \quad \forall q \in M.$$

Moreover, for all $\mathbf{w} \in \mathbf{H}^i(\Omega) \cap \mathbf{X}$, $\Phi \in \mathbf{H}^i(\Omega) \cap \mathbf{W}$ ($i = 0, 1, 2$), we set

$$\|(\mathbf{w}, \Phi)\|_i = \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} (\|\mathbf{w}\|_i^2 + \|\Phi\|_i^2)^{\frac{1}{2}}.$$

We end this section by recalling the following important conclusions.

Theorem 2.1 (See Theorems 1 and 2 of Dong et al. 2014) *Suppose that R_e, R_m, S_c , and C_1 satisfy*

$$0 < \sigma = \frac{\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|\mathbf{F}\|_{-1}}{(\min\{R_e^{-1}, S_c C_1 R_m^{-1}\})^2} < 1, \tag{2.15}$$

then problem (2.2) admits a unique solution $((\mathbf{u}, \mathbf{B}), p) \in \mathbf{W}_{0n} \times M$. Moreover,

$$\|(\mathbf{u}, \mathbf{B})\|_1 \leq \|\mathbf{F}\|_{-1}. \tag{2.16}$$

Theorem 2.2 (See Theorem 1 of Zhang et al. 2014) *Set Ω is a convex polygon/polyhedron and $0 < \sigma < 1$, if $\mathbf{f}, \mathbf{g} \in \mathbf{L}^2(\Omega)$, the solution $((\mathbf{u}, \mathbf{B}), p)$ of problem (2.2) satisfies*

$$\|(\mathbf{u}, \mathbf{B})\|_2 + \|p\|_1 \leq C\|\mathbf{F}\|_0. \tag{2.17}$$

3 The stability and convergence of iterative penalty finite element method

3.1 Finite element spaces

Set $\{\tau_\mu\}$ is a family of triangulations or tetrahedrons of Ω , and τ_μ is a shape-regular partition of Ω with mesh size μ . The real parameter $\mu > 0$ takes h or H ($h \ll H$) tending to 0. The fine grid partition τ_h is taken as a mesh refinement generated from the coarse grid τ_H . Based on the regular partitions τ_h and τ_H , we can construct the conforming finite element spaces $(\mathbf{X}_h, M_h, \mathbf{W}_h)$ and $(\mathbf{X}_H, M_H, \mathbf{W}_H) \subset (\mathbf{X}_h, M_h, \mathbf{W}_h)$. Denote $\mathbf{W}_{0n}^\mu = \mathbf{X}_\mu \times \mathbf{W}_\mu$ and assume the finite element spaces $\mathbf{X}_\mu, \mathbf{W}_\mu$ and M_μ satisfy the following assumptions.

Assumption A1 There are a mapping $r_\mu \in \mathcal{L}(\mathbf{H}^2(\Omega) \cap \mathbf{V}, \mathbf{X}_\mu)$ which satisfies

$$(\nabla \cdot (\mathbf{v} - r_\mu \mathbf{v}), q) = 0, \quad \|\nabla(\mathbf{v} - r_\mu \mathbf{v})\|_0 \leq C\mu\|\mathbf{v}\|_2, \quad \forall \mathbf{v} \in \mathbf{H}^2(\Omega) \cap \mathbf{V}, \forall q \in M_\mu,$$

and an L^2 -orthogonal projection operator $\rho_\mu: M \rightarrow M_\mu$ which satisfies

$$\|q - \rho_\mu q\|_0 \leq C\mu\|q\|_1, \quad \forall q \in H^1(\Omega) \cap M,$$

and a mapping $R_\mu \in \mathcal{L}(\mathbf{H}^2(\Omega) \cap \mathbf{V}_n, \mathbf{W}_\mu)$ which satisfies

$$\begin{aligned} (\nabla \times R_\mu \Phi, \nabla \times \Psi) + (\nabla \cdot R_\mu \Phi, \nabla \cdot \Psi) &= (\nabla \times \Phi, \nabla \times \Psi) + (\nabla \cdot \Phi, \nabla \cdot \Psi) \\ &= (\nabla \times \Phi, \nabla \times \Psi), \quad \forall \Psi \in \mathbf{W}_\mu, \\ \|\Phi - R_\mu \Phi\|_0 + \mu\|\Phi - R_\mu \Phi\|_1 &\leq C\mu^2\|\Phi\|_2, \quad \forall \Phi \in \mathbf{H}^2(\Omega) \cap \mathbf{V}_n. \end{aligned}$$

Assumption A2 Assume that the bilinear form $d_0(\cdot, \cdot)$ satisfies the discrete inf-sup condition, namely, there exists a positive constant β_0 such that:

$$\sup_{(\mathbf{v}, \Psi) \in \mathbf{W}_{0n}^\mu} \frac{|d_0((\mathbf{v}, \Psi), q)|}{\|(\mathbf{v}, \Psi)\|_1} \geq \beta_0\|q\|_0, \quad \forall q \in M_\mu.$$

There are many finite element spaces satisfying Assumptions A1 and A2 with a convex polygonal or polyhedral domain Ω . In this paper we choose the stable finite element spaces that have been used traditionally for the Navier–Stokes equations to approximate velocity and pressure. Here, the mini-element is chosen to approximate the velocity and pressure, and those finite element spaces as follows:

$$\mathbf{X}_\mu = (P_{1,\mu}^b)^d \cap \mathbf{X}, \quad M_\mu = \{q_\mu \in C^0(\Omega) : q_\mu|_K \in P_1(K), \quad \forall K \in \tau_\mu\},$$

where

$$P_{1,\mu}^b = \{v_\mu \in C^0(\Omega) : v_\mu|_K \in P_1(K) \oplus \text{span}\{\hat{b}\}, \quad \forall K \in \tau_\mu\},$$

$P_1(K)$ is defined as the space of polynomials of degree (the degree ≤ 1 on K), and \hat{b} is a bubble function. For the magnetic field approximation space \mathbf{W}_μ , there is unrestricted. For the sake of convenience, we choose the same finite element space for the magnetic field space as the one for velocity field, i.e., we use $\mathbf{W}_\mu = (P_{1,\mu}^b)^d \cap \mathbf{W}$ to approximate the magnetic field.

Now we define the discrete form of the divergence-free space \mathbf{V} as:

$$\mathbf{V}_\mu = \{ \mathbf{v} \in \mathbf{X}_\mu : d_0((\mathbf{v}, \Psi), q) = 0, \forall q \in M_\mu, \forall \Psi \in \mathbf{W}_\mu \}.$$

Introduce two L^2 -orthogonal projectors $P_\mu : \mathbf{L}^2(\Omega) \rightarrow \mathbf{V}_\mu$ and $R_{0\mu} : \mathbf{L}^2(\Omega) \rightarrow \mathbf{W}_\mu$. Define the discrete Stokes operator $A_{1\mu} = -P_\mu \Delta_\mu$, where Δ_μ is defined by (see [Sermene and Temam 1983](#))

$$-(\Delta_\mu \mathbf{u}_\mu, \mathbf{v}_\mu) = (\nabla \mathbf{u}_\mu, \nabla \mathbf{v}_\mu), \quad \forall \mathbf{u}_\mu, \mathbf{v}_\mu \in \mathbf{X}_\mu,$$

and its corresponding discrete norm is $\|\mathbf{v}_\mu\|_{j,\mu} = \|A_{1\mu}^{\frac{j}{2}} \mathbf{v}_\mu\|_0$ with the order $j \in R$, in which

$$\|\mathbf{v}_\mu\|_{1,\mu} = \|\nabla \mathbf{v}_\mu\|_0, \quad \|\mathbf{v}_\mu\|_{2,\mu} = \|A_{1\mu} \mathbf{v}_\mu\|_0, \quad \forall \mathbf{v}_\mu \in \mathbf{V}_\mu.$$

Similarly, define the discrete operator $A_{2\mu} \mathbf{B}_\mu = R_{0\mu}(\nabla_\mu \times \nabla \times \mathbf{B}_\mu + \nabla_\mu \nabla \cdot \mathbf{B}_\mu) \in \mathbf{W}_\mu$ as follows (see [He 2015](#); [Sermene and Temam 1983](#))

$$(A_{2\mu} \mathbf{B}_\mu, \Psi) = (A_{2\mu}^{\frac{1}{2}} \mathbf{B}_\mu, A_{2\mu}^{\frac{1}{2}} \Psi) = (\nabla \times \mathbf{B}_\mu, \nabla \times \Psi) + (\nabla \cdot \mathbf{B}_\mu, \nabla \cdot \Psi), \quad \forall \mathbf{B}_\mu, \Psi \in \mathbf{W}_\mu,$$

and its corresponding discrete norm is $\|\mathbf{B}_\mu\|_{j,\mu} = \|A_{2\mu}^{\frac{j}{2}} \mathbf{B}_\mu\|_0$ with the order $j \in R$, in which

$$\begin{aligned} \|\mathbf{B}_\mu\|_{1,\mu}^2 &= \|A_{2\mu}^{\frac{1}{2}} \mathbf{B}_\mu\|_0^2 = \|\nabla \times \mathbf{B}_\mu\|_0^2 + \|\nabla \cdot \mathbf{B}_\mu\|_0^2, \\ \|\mathbf{B}_\mu\|_{2,\mu} &= \|\nabla_\mu \times \nabla \times \mathbf{B}_\mu + \nabla_\mu \nabla \cdot \mathbf{B}_\mu\|_0. \end{aligned}$$

Moreover, we also introduce some discrete estimates as follows (see [Adams 1975](#); [He 2003, 2015](#))

$$\|\nabla \mathbf{v}_\mu\|_{\mathbf{L}^3} + \|\mathbf{v}_\mu\|_{\mathbf{L}^\infty} \leq C \|\nabla \mathbf{v}_\mu\|_0^{\frac{1}{2}} \|A_{1\mu} \mathbf{v}_\mu\|_0^{\frac{1}{2}}, \quad \|\nabla \mathbf{v}_\mu\|_{\mathbf{L}^6} \leq C \|A_{1\mu} \mathbf{v}_\mu\|_0, \quad \forall \mathbf{v}_\mu \in \mathbf{V}_\mu.$$

The Galerkin FEM for problem (2.2) reads as: find $((\mathbf{u}_\mu, \mathbf{B}_\mu), p_\mu) \in \mathbf{W}_{0n}^\mu \times M_\mu$ such that

$$\begin{aligned} A_0((\mathbf{u}_\mu, \mathbf{B}_\mu), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_\mu, \mathbf{B}_\mu), (\mathbf{u}_\mu, \mathbf{B}_\mu), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_\mu) \\ + d_0((\mathbf{u}_\mu, \mathbf{B}_\mu), q) = (\mathbf{F}, (\mathbf{v}, \Psi)), \quad \forall ((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n}^\mu \times M_\mu. \end{aligned} \tag{3.1}$$

Using the similar argument to Theorem 2.1, we can obtain the following conclusions (see Theorems 3 and 4 of [Dong et al. 2014](#)).

Theorem 3.1 *Under the condition of (2.15) and Assumption A1, the discrete problem (3.1) admits a unique solution $((\mathbf{u}_\mu, \mathbf{B}_\mu), p_\mu) \in \mathbf{W}_{0n}^\mu \times M_\mu$, which satisfies*

$$\|(\mathbf{u}_\mu, \mathbf{B}_\mu)\|_1 \leq \|\mathbf{F}\|_{-1}. \tag{3.2}$$

Theorem 3.2 *Under the Assumptions A1 and A2 and the condition of (2.15), the solutions of problem (3.1) satisfy*

$$\|(A_{1\mu} \mathbf{u}_\mu, A_{2\mu} \mathbf{B}_\mu)\|_0 \leq C \|\mathbf{F}\|_0. \tag{3.3}$$

Furthermore, it holds

$$\|(\mathbf{u} - \mathbf{u}_\mu, \mathbf{B} - \mathbf{B}_\mu)\|_0 + \mu(\|(\mathbf{u} - \mathbf{u}_\mu, \mathbf{B} - \mathbf{B}_\mu)\|_1 + \|p - p_\mu\|_0) \leq C \mu^2 \|\mathbf{F}\|_0. \tag{3.4}$$

3.2 Penalty finite element method

The penalty FEM for problem (2.2) is as follows: find $((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), p_{\varepsilon\mu}) \in \mathbf{W}_{0n}^\mu \times M_\mu$ such that for all $((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n}^\mu \times M_\mu$

$$A_0((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), (\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_{\varepsilon\mu}) + d_0((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), q) + \varepsilon(p_{\varepsilon\mu}, q) = \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle, \tag{3.5}$$

where $0 < \varepsilon \ll 1$ is a penalty parameter. This is the standard penalty FEM for problem (2.2). Now we present the stability and convergence of the standard penalty FEM.

Theorem 3.3 *Under the condition of (2.15) and the Assumption A1, the discrete problem (3.5) admits a unique solution $((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), p_{\varepsilon\mu}) \in \mathbf{W}_{0n}^\mu \times M_\mu$, which satisfies*

$$\|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1^2 + 2\varepsilon \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon\mu}\|_0^2 \leq \|\mathbf{F}\|_{-1}^2. \tag{3.6}$$

Furthermore, we have

$$\|p_{\varepsilon\mu}\|_0 \leq C\|\mathbf{F}\|_{-1}.$$

Proof Choosing $(\mathbf{v}, \Psi) = (\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})$ and $q = p_{\varepsilon\mu}$ in (3.5), using (2.11) and (2.9) to get

$$\begin{aligned} \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1^2 + \varepsilon \|p_{\varepsilon\mu}\|_0^2 &\leq \|\mathbf{F}\|_{-1} \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1 \\ &\leq \frac{1}{2} \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1^2 + \frac{1}{2} (\min\{R_e^{-1}, S_c C_1 R_m^{-1}\})^{-1} \|\mathbf{F}\|_{-1}^2, \end{aligned}$$

thus

$$\min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1^2 + 2\varepsilon \|p_{\varepsilon\mu}\|_0^2 \leq (\min\{R_e^{-1}, S_c C_1 R_m^{-1}\})^{-1} \|\mathbf{F}\|_{-1}^2. \tag{3.7}$$

On the other hand, taking $q = 0$ in (3.5), applying (2.10) and (2.12) to obtain

$$\begin{aligned} \beta_0 \|p_{\varepsilon\mu}\|_0 &\leq \max\{R_e^{-1}, (2+d)S_c R_m^{-1}\} \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1 \\ &\quad + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1^2 + \|\mathbf{F}\|_{-1}. \end{aligned}$$

With the help of (3.7), we have

$$\begin{aligned} \beta_0 \|p_{\varepsilon\mu}\|_0 &\leq \left[\frac{\max\{R_e^{-1}, (2+d)S_c R_m^{-1}\}}{\min\{R_e^{-1}, S_c C_1 R_m^{-1}\}} + \frac{\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|\mathbf{F}\|_{-1}}{(\min\{R_e^{-1}, S_c C_1 R_m^{-1}\})^2} + 1 \right] \|\mathbf{F}\|_{-1} \\ &\leq \left[\frac{\max\{R_e^{-1}, (2+d)S_c R_m^{-1}\}}{\min\{R_e^{-1}, S_c C_1 R_m^{-1}\}} + \sigma + 1 \right] \|\mathbf{F}\|_{-1} \leq C\|\mathbf{F}\|_{-1}, \end{aligned}$$

Thus, the proof is completed.

Theorem 3.4 *Let Ω be a convex polygonal/polyhedral domain. Under the Assumptions A1, A2 and (2.15), the solution of problem (3.5) satisfies*

$$\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu})\|_1 + \|p - p_{\varepsilon\mu}\|_0 \leq C(\mu + \varepsilon).$$

Proof Subtracting (3.5) from (2.2), we obtain the following error equation

$$\begin{aligned} A_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), (\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), (\mathbf{v}, \Psi)) \\ + A_1((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p - p_{\varepsilon\mu}) \\ + d_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), q) - \varepsilon(p_{\varepsilon\mu}, q) = 0. \end{aligned} \tag{3.8}$$

Taking $(\mathbf{v}, \Psi) = (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})$ and $q = \rho_\mu p - p_{\varepsilon\mu}$ in (3.8), using (2.9) we have

$$\begin{aligned}
 & A_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})) + \varepsilon(\rho_\mu p - p_{\varepsilon\mu}, \rho_\mu p - p_{\varepsilon\mu}) \\
 & + A_1((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), (\mathbf{u}, \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})) \\
 & = A_0((r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})) + \varepsilon(\rho_\mu p, \rho_\mu p - p_{\varepsilon\mu}) \\
 & + A_1((r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (\mathbf{u}, \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})) \\
 & + A_1((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), (r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})) \\
 & + d_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), p - p_{\varepsilon\mu}) - d_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), \rho_\mu p - p_{\varepsilon\mu}).
 \end{aligned} \tag{3.9}$$

Due to the Assumption A1, we get

$$A_0((r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})) = R_e^{-1} a_0(r_\mu \mathbf{u} - \mathbf{u}, r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}), \tag{3.10}$$

and

$$\begin{aligned}
 & d_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), p - p_{\varepsilon\mu}) - d_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), \rho_\mu p - p_{\varepsilon\mu}) \\
 & = d_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}), p - \rho_\mu p).
 \end{aligned} \tag{3.11}$$

Using (2.11) and (2.12) to obtain

$$\begin{aligned}
 & (\min\{R_e^{-1}, S_c C_1 R_m^{-1}\} - \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} \|(\mathbf{u}, \mathbf{B})\|_1) \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}\|_1^2 \\
 & + \varepsilon \|\rho_\mu p - p_{\varepsilon\mu}\|_0^2 \\
 & \leq (\{R_e^{-1} + \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} (\|(\mathbf{u}, \mathbf{B})\|_1 + \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1)\} \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 \\
 & + \sqrt{d} \|p - \rho_\mu p\|_0) \\
 & \times \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}\|_1 + \varepsilon \|\rho_\mu p\|_0 (\|p - \rho_\mu p\|_0 + \|p - p_{\varepsilon\mu}\|_0).
 \end{aligned} \tag{3.12}$$

Choosing $q = 0$ in (3.8), applying (2.10), (2.12) and Assumption A2, one finds

$$\begin{aligned}
 & \beta_0 \|p - p_{\varepsilon\mu}\|_0 \leq ((\max\{R_e^{-1}, S_c(2 + d) R_m^{-1}\} + \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} \|(\mathbf{u}, \mathbf{B})\|_1 \\
 & + \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1) \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu})\|_1 \\
 & \leq (\max\{R_e^{-1}, S_c(2 + d) R_m^{-1}\} + \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} \|(\mathbf{u}, \mathbf{B})\|_1 \\
 & + \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1) (\|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}\|_1 + \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1).
 \end{aligned} \tag{3.13}$$

Substituting (3.13) into (3.12), with the conditions of Theorem 2.1 and (3.6), we obtain

$$\begin{aligned}
 & (1 - \sigma) \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}\|_1^2 + \varepsilon \|\rho_\mu p - p_{\varepsilon\mu}\|_0^2 \\
 & \leq ((R_e^{-1} + 2\sigma \min\{R_e^{-1}, S_c C_1 R_m^{-1}\}) \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 + \sqrt{d} \|p - \rho_\mu p\|_0 \\
 & + \varepsilon \|p\|_0 \beta_0^{-1} (\max\{R_e^{-1}, S_c(2 + d) R_m^{-1}\} \\
 & + 2\sigma \min\{R_e^{-1}, S_c C_1 R_m^{-1}\})) \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}\|_1 \\
 & + \varepsilon \|p\|_0 (\|p - \rho_\mu p\|_0 + \beta_0^{-1} (\max\{R_e^{-1}, S_c(2 + d) R_m^{-1}\} \\
 & + 2\sigma \min\{R_e^{-1}, S_c C_1 R_m^{-1}\})) \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 \\
 & \leq \frac{1}{2} (1 - \sigma) \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}\|_1^2 \\
 & + C (\|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 + \|p - \rho_\mu p\|_0 + \varepsilon)^2 + C \varepsilon (\|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1
 \end{aligned}$$

$$\begin{aligned}
 & + \|p - \rho_\mu p\|_0) \\
 \leq & \frac{1}{2}(1 - \sigma) \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|(r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu})\|_1^2 \\
 & + C(\|(r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B})\|_1 + \|p - \rho_\mu p\|_0 + \varepsilon)^2.
 \end{aligned} \tag{3.14}$$

In virtue of the Assumption A1 we have

$$\begin{aligned}
 & \| (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}) \|_1 \\
 & \leq C(\|(r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B})\|_1 + \|p - \rho_\mu p\|_0 + \varepsilon) \leq C(\mu + \varepsilon).
 \end{aligned} \tag{3.15}$$

Applying the triangle inequality to gain

$$\begin{aligned}
 & \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu})\|_1 \\
 & \leq \| (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}) \|_1 + \| (r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}) \|_1 \leq C(\mu + \varepsilon).
 \end{aligned} \tag{3.16}$$

Combining (3.13) with (3.16), the error $\|p - p_{\varepsilon\mu}\|_0$ can be bounded by

$$\|p - p_{\varepsilon\mu}\|_0 \leq C \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}, \mathbf{B} - \mathbf{B}_{\varepsilon\mu})\|_1 \leq C(\mu + \varepsilon).$$

The proof of Theorem 3.4 is completed.

Next, we consider the relationship between $((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), p_{\varepsilon\mu})$ and $((\mathbf{u}_\mu, \mathbf{B}_\mu), p_\mu)$ as $\varepsilon \rightarrow 0$.

Lemma 3.5 *Let Ω be a convex polygonal/polyhedral domain. Under the Assumptions A1, A2 and (2.15), the solution $((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), p_{\varepsilon\mu})$ of problem (3.5) converges the solution $((\mathbf{u}_\mu, \mathbf{B}_\mu), p_\mu)$ of problem (3.1) as $\varepsilon \rightarrow 0$.*

Proof Subtracting (3.5) from (3.1), we obtain the following error equation

$$\begin{aligned}
 & A_0((\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu}), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu}), (\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu}), (\mathbf{v}, \Psi)) \\
 & + A_1((\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu}), (\mathbf{u}_\mu, \mathbf{B}_\mu), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_\mu - p_{\varepsilon\mu}) \\
 & + d_0((\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu}), q) - \varepsilon(p_{\varepsilon\mu}, q) = 0.
 \end{aligned} \tag{3.17}$$

Taking $(\mathbf{v}, \Psi) = (\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu})$ and $q = p_\mu - p_{\varepsilon\mu}$ in (3.17), using (2.9) we have

$$\begin{aligned}
 & A_0((\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu}), (\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu})) \\
 & + A_1((\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu}), (\mathbf{u}_\mu, \mathbf{B}_\mu), (\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu})) - \varepsilon(p_{\varepsilon\mu}, q) = 0.
 \end{aligned}$$

Using (2.11) and (2.12) to obtain

$$\begin{aligned}
 & (\min\{R_e^{-1}, S_c C_1 R_m^{-1}\} - \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}_\mu, \mathbf{B}_\mu)\|_1) \|(\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu})\|_1^2 \\
 & \leq \varepsilon \|p_{\varepsilon\mu}\|_0 \|p_\mu - p_{\varepsilon\mu}\|_0.
 \end{aligned} \tag{3.18}$$

Here, $\min\{R_e^{-1}, S_c C_1 R_m^{-1}\} - \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}_\mu, \mathbf{B}_\mu)\|_1 \geq \min\{R_e^{-1}, S_c C_1 R_m^{-1}\}(1 - \sigma) > 0$. Choosing $q = 0$ in (3.17), using (2.10), (2.12) and the Assumption A2, one finds

$$\begin{aligned}
 \beta_0 \|p_\mu - p_{\varepsilon\mu}\|_0 \leq & ((\max\{R_e^{-1}, S_c(2 + d)R_m^{-1}\} + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\}) \|(\mathbf{u}_\mu, \mathbf{B}_\mu)\|_1 \\
 & + \|(\mathbf{u}_{\varepsilon\mu}, \mathbf{B}_{\varepsilon\mu})\|_1) \|(\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu})\|_1.
 \end{aligned} \tag{3.19}$$

Substituting (3.18) into (3.19), using (3.2) and (3.6) to obtain

$$\beta_0 \|p_\mu - p_{\varepsilon\mu}\|_0 \leq C(\varepsilon \|p_{\varepsilon\mu}\|_0 \|p_\mu - p_{\varepsilon\mu}\|_0)^{1/2},$$

thus

$$\|p_\mu - p_{\varepsilon\mu}\|_0 \leq C\varepsilon \|p_{\varepsilon\mu}\|_0 \leq C\varepsilon^{1/2}. \tag{3.20}$$

Then, from (3.20) we know that $\|p_\mu - p_{\varepsilon\mu}\|_0 \rightarrow 0$ as $\varepsilon \rightarrow 0$.

Substituting (3.20) into (3.18) to obtain

$$\|(\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu})\|_1^2 \leq C\varepsilon \|p_{\varepsilon\mu}\|_0 \cdot \varepsilon^{1/2} \leq C\varepsilon. \tag{3.21}$$

From (3.21) we know that $\|(\mathbf{u}_\mu - \mathbf{u}_{\varepsilon\mu}, \mathbf{B}_\mu - \mathbf{B}_{\varepsilon\mu})\|_1 \rightarrow 0$ as $\varepsilon \rightarrow 0$. Thus the proof is finished.

3.3 Iterative penalty finite element method

The one-level iterative penalty FEM for problem (2.2) reads as:

Step 1 Find $((\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0), p_{\varepsilon\mu}^0) \in \mathbf{W}_{0n}^\mu \times M_\mu$ such that for all $((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n}^\mu \times M_\mu$

$$\begin{aligned} &A_0((\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0), (\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_{\varepsilon\mu}^0) \\ &+ d_0((\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0), q) + \varepsilon(p_{\varepsilon\mu}^0, q) = \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle. \end{aligned} \tag{3.22}$$

Step 2 For $k = 1, 2, 3, \dots$, find $((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), p_{\varepsilon\mu}^k) \in \mathbf{W}_{0n}^\mu \times M_\mu$ such that for all $\forall((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n}^\mu \times M_\mu$

$$\begin{aligned} &A_0((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_{\varepsilon\mu}^k) \\ &+ d_0((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), q) + \varepsilon(p_{\varepsilon\mu}^k, q) = \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle + \varepsilon(p_{\varepsilon\mu}^{k-1}, q). \end{aligned} \tag{3.23}$$

From above scheme, we can see that the initial value $((\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0), p_{\varepsilon\mu}^0)$ of the one-level iterative penalty FEM is gained by Step 1. From Theorems 3.3 and 3.4, we obtain the following conclusion.

Theorem 3.6 Under the conditions of Theorem 3.4, the solution $((\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0), p_{\varepsilon\mu}^0)$ of the problem (3.22) is unique and satisfies

$$\|(\mathbf{u}_{\varepsilon\mu}^0, \mathbf{B}_{\varepsilon\mu}^0)\|_1^2 + 2\varepsilon \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon\mu}^0\|_0^2 \leq \|\mathbf{F}\|_{-1}^2. \tag{3.24}$$

Furthermore, it holds

$$\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^0, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^0)\|_1 + \|p - p_{\varepsilon\mu}^0\|_0 \leq C(\mu + \varepsilon). \tag{3.25}$$

Now we study the stability of one-level iterative penalty FEM solution of (3.23).

Theorem 3.7 Under the conditions of Theorem 3.3, suppose that $((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), p_{\varepsilon\mu}^k) \in \mathbf{W}_{0n}^\mu \times M_\mu$ is the solution of the discrete problem (3.23), then the solution satisfies

$$\|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1^2 + \varepsilon \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon\mu}^k\|_0^2 \leq (k + 1) \|\mathbf{F}\|_{-1}^2. \tag{3.26}$$

Proof Taking $(\mathbf{v}, \Psi) = (\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)$ and $q = p_{\varepsilon\mu}^k$ in (3.23), using (2.9), (2.11) and (2.12) to find

$$\begin{aligned} &\min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1^2 + \varepsilon \|p_{\varepsilon\mu}^k\|_0^2 \leq \|\mathbf{F}\|_{-1} \|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1 + \varepsilon \|p_{\varepsilon\mu}^{k-1}\|_0 \|p_{\varepsilon\mu}^k\|_0 \\ &\leq \frac{1}{2} \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1^2 + \frac{1}{2} (\min\{R_e^{-1}, S_c C_1 R_m^{-1}\})^{-1} \|\mathbf{F}\|_{-1}^2 \\ &+ \frac{1}{2} \varepsilon \|p_{\varepsilon\mu}^{k-1}\|_0^2 + \frac{1}{2} \varepsilon \|p_{\varepsilon\mu}^k\|_0^2. \end{aligned}$$

Thanks to (3.24), it yields

$$\begin{aligned} & \|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1^2 + \varepsilon \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon\mu}^k\|_0^2 \\ & \leq \|\mathbf{F}\|_{-1}^2 + \varepsilon \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon\mu}^{k-1}\|_0^2 \\ & \leq k\|\mathbf{F}\|_{-1}^2 + \varepsilon \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon\mu}^0\|_0^2 \leq (k + 1)\|\mathbf{F}\|_{-1}^2, \end{aligned}$$

which implies (3.26). The proof of Theorem 3.7 is completed.

Next, we present the convergence of one-level iterative penalty FEM.

Theorem 3.8 *Under the conditions of Theorem 3.4, the solution of problem (3.23) satisfies*

$$\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 + \|p - p_{\varepsilon\mu}^k\|_0 \leq C(\mu + \varepsilon^{k+1}). \tag{3.27}$$

Proof From Theorem 3.6, we know if $k = 0$ (3.27) holds. Then we assume that (3.27) holds for $k - 1$.

From (3.23) and (2.2), we obtain

$$\begin{aligned} & A_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{v}, \Psi)) \\ & + A_1((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p - p_{\varepsilon\mu}^k) \\ & + d_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), q) - \varepsilon(p_{\varepsilon\mu}^k, q) + \varepsilon(p_{\varepsilon\mu}^{k-1}, q) = 0. \end{aligned} \tag{3.28}$$

Taking $(\mathbf{v}, \Psi) = (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)$ and $q = \rho_\mu p - p_{\varepsilon\mu}^k$ in (3.28), using (2.9) to get

$$\begin{aligned} & A_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)) \\ & + A_1((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{u}, \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)) \\ & + \varepsilon(\rho_\mu p - p_{\varepsilon\mu}^k, \rho_\mu p - p_{\varepsilon\mu}^k) \\ & = A_0((r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)) \\ & + A_1((r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (\mathbf{u}, \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)) \\ & + A_1((\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k), (r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)) \\ & + d_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), p - p_{\varepsilon\mu}^k) - d_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), \rho_\mu p - p_{\varepsilon\mu}^k) \\ & + \varepsilon(\rho_\mu p - p_{\varepsilon\mu}^{k-1}, \rho_\mu p - p_{\varepsilon\mu}^k). \end{aligned}$$

With the Assumption A1, we obtain

$$A_0((r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}), (r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)) = R_e^{-1} a_0(r_\mu \mathbf{u} - \mathbf{u}, r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k),$$

and

$$\begin{aligned} & d_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), p - p_{\varepsilon\mu}^k) - d_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), \rho_\mu p - p_{\varepsilon\mu}^k) \\ & = d_0((r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), p - \rho_\mu p). \end{aligned}$$

Using (2.11) and (2.12), one finds

$$\begin{aligned}
 & (\min\{R_e^{-1}, S_c C_1 R_m^{-1}\} - \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}, \mathbf{B})\|_1) \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1^2 \\
 & \quad + \varepsilon \|\rho_\mu p - p_{\varepsilon\mu}^k\|_0^2 \\
 & \leq (\{R_e^{-1} + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} (\|(\mathbf{u}, \mathbf{B})\|_1 + \|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1)\} \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 \\
 & \quad + \sqrt{d} \|p - \rho_\mu p\|_0) \\
 & \quad \times \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1 + \varepsilon (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^{k-1}\|_0) \\
 & \quad \times (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^k\|_0). \tag{3.29}
 \end{aligned}$$

Choosing $q = 0$ in (3.28) and combining (2.10), (2.12) and the Assumption A2 to get

$$\begin{aligned}
 \beta_0 \|p - p_{\varepsilon\mu}^k\|_0 & \leq ((\max\{R_e^{-1}, S_c(2 + d)R_m^{-1}\} + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} (\|(\mathbf{u}, \mathbf{B})\|_1 \\
 & \quad + \|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1)) \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 \\
 & \leq (\max\{R_e^{-1}, S_c(2 + d)R_m^{-1}\} + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} (\|(\mathbf{u}, \mathbf{B})\|_1 \\
 & \quad + \|(\mathbf{u}_{\varepsilon\mu}^k, \mathbf{B}_{\varepsilon\mu}^k)\|_1)) (\|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1 + \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1). \tag{3.30}
 \end{aligned}$$

Substituting (3.30) into (3.29), and using (2.16) and (3.26) to gain

$$\begin{aligned}
 & (1 - \sigma) \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1^2 + \varepsilon \|\rho_\mu p - p_{\varepsilon\mu}^k\|_0^2 \\
 & \leq ((R_e^{-1} + 2\sigma \min\{R_e^{-1}, S_c C_1 R_m^{-1}\}) \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 + \sqrt{d} \|p - \rho_\mu p\|_0 \\
 & \quad + \varepsilon (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^{k-1}\|_0) \beta_0^{-1} (\max\{R_e^{-1}, S_c(2 + d)R_m^{-1}\} \\
 & \quad + C \min\{R_e^{-1}, S_c C_1 R_m^{-1}\})) \\
 & \quad \times \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1 + \varepsilon (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^{k-1}\|_0) (\|p - \rho_\mu p\|_0 \\
 & \quad + \beta_0^{-1} (\max\{R_e^{-1}, S_c(2 + d)R_m^{-1}\} + C \min\{R_e^{-1}, S_c C_1 R_m^{-1}\}) \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1) \\
 & \leq \frac{1}{2} (1 - \sigma) \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1^2 \\
 & \quad + C (\|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 + \|p - \rho_\mu p\|_0 + \varepsilon (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^{k-1}\|_0))^2 \\
 & \quad + C \varepsilon (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^{k-1}\|_0) (\|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 + \|p - \rho_\mu p\|_0) \\
 & \leq \frac{1}{2} (1 - \sigma) \min\{R_e^{-1}, S_c C_1 R_m^{-1}\} \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1^2 \\
 & \quad + C (\|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 + \|p - \rho_\mu p\|_0 + \varepsilon (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^{k-1}\|_0))^2.
 \end{aligned}$$

Using the Assumption A1 to get

$$\begin{aligned}
 \| \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1 \| & \leq C (\|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 + \|p - \rho_\mu p\|_0 \\
 & \quad + \varepsilon (\|\rho_\mu p - p\|_0 + \|p - p_{\varepsilon\mu}^{k-1}\|_0)) \leq C(\mu + \varepsilon^{k+1}).
 \end{aligned}$$

From the triangle inequality we gain

$$\begin{aligned}
 \| \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 \| & \leq \| \|r_\mu \mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, R_\mu \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k\|_1 \| + \| \|r_\mu \mathbf{u} - \mathbf{u}, R_\mu \mathbf{B} - \mathbf{B}\|_1 \| \\
 & \leq C(\mu + \varepsilon^{k+1}). \tag{3.31}
 \end{aligned}$$

Combining (3.30) with (3.31), the error estimate $\|p - p_{\varepsilon\mu}^k\|_0$ can be bounded by

$$\|p - p_{\varepsilon\mu}^k\|_0 \leq C \| \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 \| \leq C(\mu + \varepsilon^{k+1}). \tag{3.32}$$

Thus, the proof of Theorem 3.8 is completed.

Next, we present the L^2 error estimate $\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0$. To achieve this aim, we consider the following dual problem: find $((\mathbf{w}, \Phi), s) \in \mathbf{W}_{0n} \times M$ such that

$$\begin{aligned} & A_0((\mathbf{v}, \Psi), (\mathbf{w}, \Phi)) + A_1((\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi), (\mathbf{w}, \Phi)) + A_1((\mathbf{v}, \Psi), (\mathbf{u}, \mathbf{B}), (\mathbf{w}, \Phi)) \\ & - d_0((\mathbf{v}, \Psi), s) + d_0((\mathbf{w}, \Phi), q) \\ & = ((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{v}, \Psi)), \quad \forall ((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n} \times M. \end{aligned} \tag{3.33}$$

If the solution of problem (3.33) satisfies $\mathbf{w} \in \mathbf{H}^2(\Omega) \cap \mathbf{X}$, $\Phi \in \mathbf{H}^2(\Omega) \cap \mathbf{W}$, then we have (see Gunzburger et al. 1991)

$$\|(\mathbf{w}, \Phi)\|_2 + \|s\|_1 \leq C \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0. \tag{3.34}$$

Theorem 3.9 Under the conditions of Theorem 3.4, the solution of problem (3.23) satisfies

$$\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0 \leq C(\mu^2 + \mu\varepsilon + \varepsilon^{k+1}). \tag{3.35}$$

Proof Choosing $(\mathbf{v}, \Psi) = (r_\mu \mathbf{w}, R_\mu \Phi)$ and $q = -\rho_\mu s$ in (3.28), subtracting it from (3.33) with $(\mathbf{v}, \Psi) = (\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)$ and $q = p_{\varepsilon\mu}^k - p$, we obtain

$$\begin{aligned} & \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0^2 = A_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)) \\ & + A_1((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{u}, \mathbf{B}), (\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)) \\ & + A_1((\mathbf{u}, \mathbf{B}), (\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)) \\ & - A_1((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)) \\ & + A_1((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), (\mathbf{w}, \Phi)) \\ & - d_0((\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi), p - p_{\varepsilon\mu}^k) - d_0((\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k), s - \rho_\mu s) \\ & - \varepsilon(p_{\varepsilon\mu}^k, \rho_\mu s) + \varepsilon(p_{\varepsilon\mu}^{k-1}, \rho_\mu s). \end{aligned}$$

Applying the conditions of Theorem 2.1, (2.10), (2.12), (3.6) and (3.34), one finds

$$\begin{aligned} & \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0^2 \\ & \leq \max\{R_e^{-1}, (2 + d)S_c R_m^{-1}\} \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 \|(\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)\|_1 \\ & + 2\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}, \mathbf{B})\|_1 \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 \|(\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)\|_1 \\ & + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1^2 (\|(\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)\|_1 + \|(\mathbf{w}, \Phi)\|_1) \\ & + \sqrt{d} \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 \|s - \rho_\mu s\|_0 + \sqrt{d} \|(\mathbf{w} - r_\mu \mathbf{w}, \Phi - R_\mu \Phi)\|_1 \|p - p_{\varepsilon\mu}^k\|_0 \\ & \leq C(\mu \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 + \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1^2) \|(\mathbf{w}, \Phi)\|_2 \\ & + C\mu (\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 \|s\|_1 + \|(\mathbf{w}, \Phi)\|_2 \|p - p_{\varepsilon\mu}^k\|_0) \\ & + \varepsilon \|p - p_{\varepsilon\mu}^k\|_0 \|s\|_0 + \varepsilon \|p - p_{\varepsilon\mu}^{k-1}\|_0 \|s\|_0 \\ & \leq C(\mu (\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 + \|p - p_{\varepsilon\mu}^k\|_0) + \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1^2) \\ & + \varepsilon \|p - p_{\varepsilon\mu}^k\|_0 + \varepsilon \|p - p_{\varepsilon\mu}^{k-1}\|_0 \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0. \end{aligned}$$

Thanks to the Theorem 3.8, we have

$$\begin{aligned} \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_0 &\leq C \left(\mu(\|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1 + \|p - p_{\varepsilon\mu}^k\|_0) \right. \\ &\quad \left. + \|(\mathbf{u} - \mathbf{u}_{\varepsilon\mu}^k, \mathbf{B} - \mathbf{B}_{\varepsilon\mu}^k)\|_1^2 + \varepsilon\|p - p_{\varepsilon\mu}^k\|_0 + \varepsilon\|p - p_{\varepsilon\mu}^{k-1}\|_0 \right) \\ &\leq C(\mu(\mu + \varepsilon^{k+1}) + (\mu + \varepsilon^{k+1})^2 + \varepsilon(\mu + \varepsilon^k)) \leq C(\mu^2 + \mu\varepsilon + \varepsilon^{k+1}). \end{aligned}$$

As a consequence, the desired result is obtained.

4 Two-level iterative penalty finite element method

In this section, we consider the stability and convergence of two-level iterative penalty FEM for the stationary incompressible MHD problem.

The two-level iterative penalty FEM based on Stokes iteration can be described as follows.

Step 1 Find $((\mathbf{u}_{\varepsilon H}^0, \mathbf{B}_{\varepsilon H}^0), p_{\varepsilon H}^0) \in \mathbf{W}_{0n}^H \times M_H$ such that for all $((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n}^H \times M_H$

$$\begin{aligned} A_0((\mathbf{u}_{\varepsilon H}^0, \mathbf{B}_{\varepsilon H}^0), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon H}^0, \mathbf{B}_{\varepsilon H}^0), (\mathbf{u}_{\varepsilon H}^0, \mathbf{B}_{\varepsilon H}^0), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_{\varepsilon H}^0) \\ + d_0((\mathbf{u}_{\varepsilon H}^0, \mathbf{B}_{\varepsilon H}^0), q) + \varepsilon(p_{\varepsilon H}^0, q) = \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle. \end{aligned} \tag{4.1}$$

Step 2 For $n = 1, 2, 3, \dots, k$, find $((\mathbf{u}_{\varepsilon H}^n, \mathbf{B}_{\varepsilon H}^n), p_{\varepsilon H}^n) \in \mathbf{W}_{0n}^H \times M_H$ such that for all $((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n}^H \times M_H$

$$\begin{aligned} A_0((\mathbf{u}_{\varepsilon H}^n, \mathbf{B}_{\varepsilon H}^n), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon H}^n, \mathbf{B}_{\varepsilon H}^n), (\mathbf{u}_{\varepsilon H}^n, \mathbf{B}_{\varepsilon H}^n), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_{\varepsilon H}^n) \\ + d_0((\mathbf{u}_{\varepsilon H}^n, \mathbf{B}_{\varepsilon H}^n), q) + \varepsilon(p_{\varepsilon H}^n, q) = \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle + \varepsilon(p_{\varepsilon H}^{n-1}, q). \end{aligned} \tag{4.2}$$

In step 3, we solve a Stokes iterative MHD problem on fine mesh.

Step 3 Find $((\mathbf{u}_{\varepsilon}^h, \mathbf{B}_{\varepsilon}^h), p_{\varepsilon}^h) \in \mathbf{W}_{0n}^h \times M_h$ such that for any $((\mathbf{v}, \Psi), q) \in \mathbf{W}_{0n}^h \times M_h$

$$\begin{aligned} A_0((\mathbf{u}_{\varepsilon}^h, \mathbf{B}_{\varepsilon}^h), (\mathbf{v}, \Psi)) + A_1((\mathbf{u}_{\varepsilon H}^k, \mathbf{B}_{\varepsilon H}^k), (\mathbf{u}_{\varepsilon}^h, \mathbf{B}_{\varepsilon}^h), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p_{\varepsilon}^h) \\ + d_0((\mathbf{u}_{\varepsilon}^h, \mathbf{B}_{\varepsilon}^h), q) + \varepsilon(p_{\varepsilon}^h, q) = \langle \mathbf{F}, (\mathbf{v}, \Psi) \rangle + \varepsilon(p_{\varepsilon H}^k, q). \end{aligned} \tag{4.3}$$

Remark 4.1 In our two-level iterative penalty FEM, we adopt the Stokes iteration to treat the nonlinear terms, other iterative schemes, such as the Newton iteration and Oseen iteration, can also be used to treat the nonlinear terms. Here, we omit the analysis of these iterative schemes due to the similar proofs.

Now we present the stability of the two-level iterative penalty FEM.

Theorem 4.2 *Under the conditions of Theorem 3.3, the solution $((\mathbf{u}_{\varepsilon}^h, \mathbf{B}_{\varepsilon}^h), p_{\varepsilon}^h)$ defined by scheme (4.3) satisfies*

$$\|(\mathbf{u}_{\varepsilon}^h, \mathbf{B}_{\varepsilon}^h)\|_1 + \varepsilon \min\{R_{\varepsilon}^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon}^h\|_0 \leq (k^2 + 5k + 5) \|\mathbf{F}\|_2^2. \tag{4.4}$$

where k is the number of iterative step.

Proof Choosing $(\mathbf{v}, \Psi) = (\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h)$ and $q = p_\varepsilon^h$ in (4.3), and using (2.11), (2.12), (2.15) and (3.26), we gain

$$\begin{aligned} & \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h)\|_1^2 + \varepsilon \|p_\varepsilon^h\|_0^2 \\ & \leq (\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}_{\varepsilon H}^k, \mathbf{B}_{\varepsilon H}^k)\|_1^2 + \|\mathbf{F}\|_{-1}) \|(\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h)\|_1 + \varepsilon \|p_\varepsilon^h\|_0 \|p_{\varepsilon H}^k\|_0 \\ & \leq (1 + \sigma(k + 1)) \|\mathbf{F}\|_{-1} \|(\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h)\|_1 + \varepsilon \|p_\varepsilon^h\|_0 \|p_{\varepsilon H}^k\|_0 \\ & \leq \frac{1}{2} \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|(\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h)\|_1^2 + \frac{(k + 2)^2}{2} (\min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\})^{-1} \|\mathbf{F}\|_{-1}^2 \\ & \quad + \frac{1}{2} \varepsilon \|p_\varepsilon^h\|_0^2 + \frac{1}{2} \varepsilon \|p_{\varepsilon H}^k\|_0^2. \end{aligned}$$

As a consequence one finds

$$\begin{aligned} & \|(\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h)\|_1^2 + \varepsilon \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|p_\varepsilon^h\|_0^2 \\ & \leq (k + 2)^2 \|\mathbf{F}\|_{-1}^2 + \varepsilon \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|p_{\varepsilon H}^k\|_0^2 \leq (k^2 + 5k + 5) \|\mathbf{F}\|_{-1}^2. \end{aligned}$$

Then the proof is completed.

Theorem 4.3 *Under the conditions of Theorem 3.6, the solution $((\mathbf{u}_\varepsilon^h, \mathbf{B}_\varepsilon^h), p_\varepsilon^h)$ of two-level iterative penalty FEM defined by scheme (4.3) satisfies*

$$\|(\mathbf{u} - \mathbf{u}_\varepsilon^h, \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 + \|p - p_\varepsilon^h\|_0 \leq C(h + H^2 + H\varepsilon + \varepsilon^{k+1}). \tag{4.5}$$

Proof Subtracting (4.3) from (2.2), we have

$$\begin{aligned} & A_0((\mathbf{u} - \mathbf{u}_\varepsilon^h, \mathbf{B} - \mathbf{B}_\varepsilon^h), (\mathbf{v}, \Psi)) + A_1((\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k), (\mathbf{u}, \mathbf{B}), (\mathbf{v}, \Psi)) \\ & \quad + A_1((\mathbf{u}, \mathbf{B}), (\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k), (\mathbf{v}, \Psi)) - A_1((\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k), \\ & (\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k), (\mathbf{v}, \Psi)) - d_0((\mathbf{v}, \Psi), p - p_\varepsilon^h) + d_0((\mathbf{u} - \mathbf{u}_\varepsilon^h, \mathbf{B} - \mathbf{B}_\varepsilon^h), q) \\ & \quad - \varepsilon(p_\varepsilon^h, q) + \varepsilon(p_{\varepsilon H}^k, q) = 0. \end{aligned} \tag{4.6}$$

Taking $(\mathbf{v}, \Psi) = (r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)$ and $q = \rho_h p - p_\varepsilon^h$ in (4.6), using (2.11), (2.12), (2.13) and (2.14) to gain

$$\begin{aligned} & \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1^2 + \varepsilon \|\rho_h p - p_\varepsilon^h\|_0^2 \\ & \leq (R_\varepsilon^{-1} \|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1 + C\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}, \mathbf{B})\|_2 \\ & \quad \times \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2 \\ & \quad + \sqrt{d} \|\rho_h p - p\|_0) \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 \\ & \quad + \varepsilon (\|\rho_h p - p\|_0 + \|p - p_{\varepsilon H}^k\|_0) (\|\rho_h p - p\|_0 + \|p - p_\varepsilon^h\|_0). \end{aligned} \tag{4.7}$$

Taking $q = 0$ in (4.6), thanks to (2.10), (2.12), (2.13), (2.14) and Assumption A2, one finds

$$\begin{aligned} \beta_0 \|p - p_\varepsilon^h\|_0 & \leq C\sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u}, \mathbf{B})\|_2 \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 \\ & \quad + \max\{R_\varepsilon^{-1}, (2 + d)S_c R_m^{-1}\} \|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1 \\ & \quad + \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 \\ & \quad + \sqrt{2}C_0^2 \max\{1, \sqrt{2}S_c\} \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2. \end{aligned} \tag{4.8}$$

Substituting (4.8) into (4.7) and applying (2.17) we obtain

$$\begin{aligned}
 & \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1^2 + \varepsilon \|\rho_h p - p_\varepsilon^h\|_0^2 \\
 & \leq (R_\varepsilon^{-1} \|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1 + C \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} \|(\mathbf{u}, \mathbf{B})\|_2 \\
 & \quad \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 \\
 & \quad + \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2 + \sqrt{d} \|\rho_h p - p\|_0 \\
 & \quad + \varepsilon (\|\rho_h p - p\|_0 + \|p - p_{\varepsilon H}^k\|_0) \beta_0^{-1} \max\{R_\varepsilon^{-1}, (2 + d) S_c R_m^{-1}\} \\
 & \quad \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 \\
 & \quad + \varepsilon (\|\rho_h p - p\|_0 + \|p - p_{\varepsilon H}^k\|_0) (\|\rho_h p - p\|_0 + \beta_0^{-1} (C \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} \|(\mathbf{u}, \mathbf{B})\|_2 \\
 & \quad \cdot \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 + \max\{R_\varepsilon^{-1}, (2 + d) S_c R_m^{-1}\} \|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1 \\
 & \quad + \sqrt{2} C_0^2 \max\{1, \sqrt{2} S_c\} \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2)) \\
 & \leq \frac{1}{2} \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1^2 + C (\|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1 \\
 & \quad + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2 + \|\rho_h p - p\|_0 \\
 & \quad + \varepsilon (\|\rho_h p - p\|_0 + \|p - p_{\varepsilon H}^k\|_0)^2 + C \varepsilon (\|\rho_h p - p\|_0 + \|p - p_{\varepsilon H}^k\|_0) (\|\rho_h p - p\|_0 \\
 & \quad + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2 + \|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1) \\
 & \leq \frac{1}{2} \min\{R_\varepsilon^{-1}, S_c C_1 R_m^{-1}\} \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1^2 + C (\|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1 \\
 & \quad + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2 + \|\rho_h p - p\|_0 \\
 & \quad + \varepsilon (\|\rho_h p - p\|_0 + \|p - p_{\varepsilon H}^k\|_0)^2).
 \end{aligned}$$

Using the Theorems 3.8 and 3.9 to obtain

$$\begin{aligned}
 \|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 & \leq C (\|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1 + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_0 \\
 & \quad + \|(\mathbf{u} - \mathbf{u}_{\varepsilon H}^k, \mathbf{B} - \mathbf{B}_{\varepsilon H}^k)\|_1^2 + \|\rho_h p - p\|_0 + \varepsilon (\|\rho_h p - p\|_0 + \|p - p_{\varepsilon H}^k\|_0)) \\
 & \leq C (h + (H^2 + H\varepsilon + \varepsilon^{k+1}) + (H + \varepsilon^{k+1})^2 + \varepsilon (h + H + \varepsilon^{k+1})) \\
 & \leq C (h + H^2 + H\varepsilon + \varepsilon^{k+1}).
 \end{aligned}$$

By the triangle inequality and the Assumption A1, it holds

$$\begin{aligned}
 \|(\mathbf{u} - \mathbf{u}_\varepsilon^h, \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 & \leq C (\|(r_h \mathbf{u} - \mathbf{u}_\varepsilon^h, R_h \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 + \|(r_h \mathbf{u} - \mathbf{u}, R_h \mathbf{B} - \mathbf{B})\|_1). \\
 & \leq C (h + H^2 + H\varepsilon + \varepsilon^{k+1}).
 \end{aligned} \tag{4.9}$$

From (4.8) and (4.9), we have

$$\|p - p_\varepsilon^h\|_0 \leq C (h + H^2 + H\varepsilon + \varepsilon^{k+1}). \tag{4.10}$$

We finish the proof by combining (4.9) with (4.10).

Remark 4.4 If we take $\varepsilon = \mathcal{O}(H)$ and $h = \mathcal{O}(H^2)$ for the two-level iterative penalty FEM, we can get the same order of convergence rate as the standard Galerkin FEM, namely, it holds

$$\|(\mathbf{u} - \mathbf{u}_\varepsilon^h, \mathbf{B} - \mathbf{B}_\varepsilon^h)\|_1 + \|p - p_\varepsilon^h\|_0 \leq Ch.$$

5 Numerical analysis

In this section, we present some numerical results of one-level and two-level iterative penalty FEMs for incompressible MHD equations. The software FreeFem++ is used in this numerical experiments (see Hecht et al. 2015). The UMFPACK routine is applied to solve the linear systems arising from the discrete algebraic equations. The mesh consists of triangular elements that are obtained by dividing Ω into subsquares of equal size and then drawing the diagonal in each sub-square. The (P_1b, P_1, P_1b) finite element pair is used and the iterative tolerance 10^{-5} is adopted in all numerical tests.

The example is quoted from Tao and Zhang (2015). The steady incompressible MHD equations are defined on a convex domain $\Omega = [0, 1]^2$. The boundary and initial conditions and right-hand side functions \mathbf{f} and \mathbf{g} are selected such that the exact solutions are given by

$$u_1 = x^2(x - 1)^2y(y - 1)(2y - 1); \quad u_2 = -y^2(y - 1)^2x(x - 1)(2x - 1);$$

$$p = (2x - 1)(2y - 1); \quad B_1 = \sin(\pi x) \cos(\pi y); \quad B_2 = -\sin(\pi y) \cos(\pi x);$$

where the components of \mathbf{u} and \mathbf{B} are denoted by (u_1, u_2) and (B_1, B_2) for convenience. Firstly, we choose the parameters $R_e = R_m = S_c = 1$ and $\varepsilon = 0.001$. In all numerical tests, we use several mesh pairs $1/h = 9, 16, 25, 36, 49, 64, 81, 100$ and $H = h^{\frac{1}{2}}$. Comparison of relative errors with different iterations are shown in Tables 1 and 2 for one-level and two-level iterative penalty FEMs respectively. Then we show the relative errors between the exact solution and the numerical solutions obtained from one-level and two-level iterative penalty FEMs in Tables 3 and 4. As observed from Tables 3 and 4, the errors $\frac{\|\mathbf{u}-\mathbf{u}_h\|_1}{\|\mathbf{u}\|_1}$, $\frac{\|\mathbf{B}-\mathbf{B}_h\|_1}{\|\mathbf{B}\|_1}$, $\frac{\|\mathbf{u}-\mathbf{u}_h\|_0}{\|\mathbf{u}\|_0}$ and $\frac{\|\mathbf{B}-\mathbf{B}_h\|_0}{\|\mathbf{B}\|_0}$ become smaller and smaller as the mesh is refined. In all tables, the symbol ‘‘Iteration’’ denotes the number of iteration in Step 2 of corresponding method. From these tables, the observations and conclusions are obtained as follows:

- Based on Tables 1 and 2, the errors of the velocity and magnetic in \mathbf{H}^1 - and \mathbf{L}^2 -norms become smaller as the iteration increasing in both one-level and two-level iterative penalty

Table 1 One-level iterative penalty FEM for incompressible MHD problem

$\frac{\ \mathbf{u}-\mathbf{u}_h\ _0}{\ \mathbf{u}\ _0}$	$h = 1/6$	$h = 1/12$	$h = 1/18$	$\frac{\ \mathbf{u}-\mathbf{u}_h\ _1}{\ \mathbf{u}\ _1}$	$h = 1/6$	$h = 1/12$	$h = 1/18$
$k = 1$	0.211875	0.0518668	0.0227872	$k = 1$	0.621663	0.243481	0.15372
$k = 2$	0.202109	0.0515975	0.0227779	$k = 2$	0.516145	0.232147	0.150471
$k = 3$	0.202108	0.0515977	0.0227781	$k = 3$	0.516145	0.232147	0.150471
$k = 4$	0.202108	0.0515977	0.0227781	$k = 4$	0.516145	0.232147	0.150471
$k = 5$	0.202108	0.0515977	0.0227781	$k = 5$	0.516145	0.232147	0.150471
$\frac{\ \mathbf{B}-\mathbf{B}_h\ _0}{\ \mathbf{B}\ _0}$	$h = 1/6$	$h = 1/12$	$h = 1/18$	$\frac{\ \mathbf{B}-\mathbf{B}_h\ _1}{\ \mathbf{B}\ _1}$	$h = 1/6$	$h = 1/12$	$h = 1/18$
$k = 1$	0.0673653	0.0174816	0.00782663	$k = 1$	0.243099	0.123492	0.0825815
$k = 2$	0.0673654	0.0174816	0.00782657	$k = 2$	0.243099	0.123492	0.0825815
$k = 3$	0.0673654	0.0174816	0.00782657	$k = 3$	0.243099	0.123492	0.0825815
$k = 4$	0.0673654	0.0174816	0.00782657	$k = 4$	0.243099	0.123492	0.0825815
$k = 5$	0.0673654	0.0174816	0.00782657	$k = 5$	0.243099	0.123492	0.0825815

Table 2 Two-level iterative penalty FEM for incompressible MHD problem with $H = h^{1/2}$

$\frac{\ \mathbf{u}-\mathbf{u}_h\ _0}{\ \mathbf{u}\ _0}$	$h = 1/9$	$h = 1/16$	$h = 1/25$	$\frac{\ \mathbf{u}-\mathbf{u}_h\ _1}{\ \mathbf{u}\ _1}$	$h = 1/9$	$h = 1/16$	$h = 1/25$
$k = 1$	0.115708	0.0611698	0.0397726	$k = 1$	0.375119	0.209218	0.13314
$k = 2$	0.0917558	0.0288856	0.0117347	$k = 2$	0.319636	0.170385	0.106938
$k = 3$	0.0917622	0.0288912	0.0117391	$k = 3$	0.319636	0.170385	0.106938
$k = 4$	0.0917622	0.0288912	0.0117391	$k = 4$	0.319636	0.170385	0.106938
$k = 5$	0.0917622	0.0288912	0.0117391	$k = 5$	0.319636	0.170385	0.106938

$\frac{\ \mathbf{B}-\mathbf{B}_h\ _0}{\ \mathbf{B}\ _0}$	$h = 1/9$	$h = 1/16$	$h = 1/25$	$\frac{\ \mathbf{B}-\mathbf{B}_h\ _1}{\ \mathbf{B}\ _1}$	$h = 1/9$	$h = 1/16$	$h = 1/25$
$k = 1$	0.0307716	0.00989297	0.00407113	$k = 1$	0.163963	0.0928432	0.05953
$k = 2$	0.0307694	0.00989001	0.00406892	$k = 2$	0.163963	0.0928431	0.05953
$k = 3$	0.0307694	0.00989001	0.00406892	$k = 3$	0.163963	0.0928431	0.05953
$k = 4$	0.0307694	0.00989001	0.00406892	$k = 4$	0.163963	0.0928431	0.05953
$k = 5$	0.0307694	0.00989001	0.00406892	$k = 5$	0.163963	0.0928431	0.05953

Table 3 One-level iterative penalty FEM for incompressible MHD problem

$1/h$	$\frac{\ \mathbf{u}-\mathbf{u}_h\ _0}{\ \mathbf{u}\ _0}$	Rate	$\frac{\ \mathbf{u}-\mathbf{u}_h\ _1}{\ \mathbf{u}\ _1}$	Rate	$\frac{\ p-p_h\ _0}{\ p\ _0}$	Rate
9	0.0917622		0.319636		0.0491044	
16	0.0288912	2.00858	0.170385	1.09343	0.0161331	1.93456
25	0.0117391	2.01801	0.106938	1.04375	0.00680877	1.93297
36	0.00563024	2.01506	0.0736314	1.02341	0.00338634	1.91547
49	0.00302794	2.01189	0.0538613	1.01413	0.0018876	1.8957
64	0.00177046	2.00943	0.0411348	1.00934	0.00114322	1.87768
81	0.00110332	2.00755	0.032451	1.00661	0.000737113	1.863
100	0.000722968	2.00605	0.0262581	1.00491	0.000498938	1.85202

$1/h$	$\frac{\ \mathbf{B}-\mathbf{B}_h\ _0}{\ \mathbf{B}\ _0}$	Rate	$\frac{\ \mathbf{B}-\mathbf{B}_h\ _1}{\ \mathbf{B}\ _1}$	Rate	CPU(S)	Iteration
9	0.0307694		0.163963		0.764	2
16	0.00989001	1.97266	0.0928431	0.988471	2.013	2
25	0.00406892	1.99008	0.05953	0.995842	4.868	2
36	0.0019654	1.99559	0.0413681	0.998156	10.124	2
49	0.00106161	1.99776	0.0304017	0.999062	18.689	2
64	0.000622505	1.99874	0.0232795	0.999473	31.934	2
81	0.000388696	1.99924	0.0183951	0.999682	53.555	2
100	0.00025505	1.99951	0.0149007	0.999797	79.638	2

methods. Especially, when $k = 2$ the results is as good as k takes 3, 4, and 5. Thus we choose the iteration $k = 2$ in following numerical tests.

- From Table 3, we can see that the optimal numerical convergence orders of one-level iterative penalty FEM are agreed with the ones predicted by the theoretical analysis in

Table 4 Two-level iterative penalty FEM for incompressible MHD problem

$1/H$	$1/h$	$\frac{\ \mathbf{u}-\mathbf{u}_h\ _0}{\ \mathbf{u}\ _0}$	Rate	$\frac{\ \mathbf{u}-\mathbf{u}_h\ _1}{\ \mathbf{u}\ _1}$	Rate	$\frac{\ p-p_h\ _0}{\ p\ _0}$	Rate
3	9	0.0917558		0.319636		0.049104	
4	16	0.0288856	2.00879	0.170385	1.09344	0.016135	1.93434
5	25	0.0117347	2.01841	0.106938	1.04375	0.00680994	1.93285
6	36	0.00562686	2.01567	0.0736312	1.02341	0.00338713	1.9153
7	49	0.00302532	2.01275	0.0538612	1.01413	0.00188816	1.89549
8	64	0.00176839	2.01056	0.0411347	1.00934	0.00114363	1.87743
9	81	0.00110166	2.009	0.032451	1.00661	0.000737433	1.8627
10	100	0.000721607	2.00784	0.0262581	1.00491	0.000499192	1.85167
$1/H$	$1/h$	$\frac{\ \mathbf{B}-\mathbf{B}_h\ _0}{\ \mathbf{B}\ _0}$	Rate	$\frac{\ \mathbf{B}-\mathbf{B}_h\ _1}{\ \mathbf{B}\ _1}$	Rate	CPU(S)	Iteration
3	9	0.0307694		0.163963		0.377	2
4	16	0.00989001	1.97266	0.0928431	0.988471	0.998	2
5	25	0.00406892	1.99008	0.05953	0.995842	2.372	2
6	36	0.0019654	1.99559	0.0413681	0.998156	4.576	2
7	49	0.00106161	1.99776	0.0304017	0.999062	8.271	2
8	64	0.000622505	1.99874	0.0232795	0.999473	13.952	2
9	81	0.000388696	1.99924	0.0183951	0.999682	22.624	2
10	100	0.00025505	1.99951	0.0149007	0.999797	34.298	2

Theorems 3.8 and 3.9, namely, $\mathcal{O}(h)$ for velocity and magnetic in \mathbf{H}^1 -norm and pressure in L^2 -norm, and $\mathcal{O}(h^2)$ for velocity and magnetic in \mathbf{L}^2 -norm.

- From Table 4, two-level iterative penalty FEM can achieve the optimal numerical convergence orders of $\mathcal{O}(h)$ for velocity and magnetic in \mathbf{H}^1 -norm and pressure in L^2 -norm, as proven in Theorem 4.3. Furthermore, we can find that two-level iterative penalty FEM can reach the optimal orders of $\mathcal{O}(h^2)$ for velocity and magnetic in \mathbf{L}^2 -norm.
- By comparing the Tables 3 and 4, we can see that two-level iterative penalty FEM significantly takes the least CPU time than the one-level iterative penalty FEM with the same approximation results.

6 Conclusion

In this paper, we present the theoretical analysis of the one-level and two-level iterative penalty FEMs for the steady incompressible MHD problem. The stability and error estimates of these numerical methods are obtained. Numerical experiments are made to show that the one-level and two-level iterative penalty FEMs are valid for solving the incompressible MHD problem, and the numerical results are consistent with the theoretical analysis. Moreover, in our further works we will consider the extensions of the Stokes iteration on fine mesh to other linearization methods, such as the Oseen and Newton iterations, combining the present methods with some stabilization techniques likes subgrid method or variational multiscale method, and solving large Reynolds number MHD problem.

References

- Adams RA (1975) Sobolev space. Academic Press, New York
- An R, Shi F (2015) Two-level iteration penalty methods for the incompressible flows. *Appl Math Model* 39:630–641
- Dai X (2007) Finite element approximation of the pure Neumann problem using the iterative penalty method. *Appl Math Comput* 186:1367–1373
- Discacciati M (2008) Numerical approximation of a steady MHD problem. Springer, New York
- Dong XJ, He YN, Zhang Y (2014) Convergence analysis of three finite element iterative methods for the 2D/3D stationary incompressible magnetohydrodynamics. *Comput Methods Appl Mech Eng* 276:287–311
- Girault V, Lions JL (2001) Two-grid finite element scheme for the steady Navier–Stokes equations in polyhedra. *Port Math* 58:25–57
- Girault V, Raviart PA (1986) Finite element approximation of Navier–Stokes equations. Springer, Berlin
- Gunzburger MD (1989) Iterated penalty methods for the Stokes and Navier–Stokes equations. *Finite Elem Anal Fluids* 1:1040–1045
- Gunzburger MD, Meir AJ, Peterson JS (1991) On the existence, uniqueness, and finite element approximation of solutions of the equations of stationary incompressible magnetohydrodynamics. *Math Comput* 56:523–563
- Gunzburger MD, Ladyzhenskaya OA, Peterson JS (2004) On the global unique solvability and initial boundary value problems for coupled modified Navier–Stokes and Maxwell equations. *J Math Fluid Mech* 6:462–482
- Hasler U, Schneebeli A, Schözau D (2004) Mixed finite element approximation of incompressible MHD problems based on weighted regularization. *Appl Numer Math* 51:19–45
- He YN (2003) Two-level method based on finite element and Crank–Nicolson extrapolation for the time-dependent Navier–Stokes equations. *SIAM J Numer Anal* 41:1263–1285
- He YN (2004) A two-level finite element Galerkin method for the nonstationary Navier–Stokes equations I: spatial discretization. *J Comput Math* 22:21–32
- He YN (2005) Optimal error estimate of the penalty finite element method for the time-dependent Navier–Stokes equations. *Math Comput* 74:1201–1216
- He YN (2015) Unconditional convergence of the Euler semi-implicit scheme for the 3D incompressible MHD equations. *IMA J Numer Anal* 35:767–801
- Hecht F, Pironneau O, Hyaric A, Ohtsuka K (2015) FreeFem++, version 3.37, 2008. <http://www.freefem.org>. Accessed 11 May 2015
- Hughes WF, Young FJ (1966) The electromagnetics of fluids. Wiley, New York
- Marion M, Xu JC (1995) Error estimates on a new nonlinear Galerkin method based on two-grid finite elements. *SIAM J Numer Anal* 32:1170–1184
- Moreau R (1990) Magneto-hydrodynamics. Kluwer Academic Publishers, Dordrecht
- Qiu HL, Mei LQ, Zhang YM (2014) Iterative penalty methods for the steady Navier–Stokes equations. *Appl Math Comput* 237:110–119
- Schözau D (2004) Mixed finite element methods for stationary incompressible magneto-hydrodynamics. *Numer Math* 96:771–800
- Sermame M, Temam R (1983) Some mathematical questions related to the MHD equations. *Commun Pure Appl Math* 36(5):635–664
- Shen J (1995) On error estimates of the penalty method for unsteady Navier–Stokes equations. *SIAM J Numer Anal* 32:386–403
- Tao ZZ, Zhang T (2015) Stability and convergence of two-level iterative methods for the stationary incompressible magnetohydrodynamics with different Reynolds numbers. *J Math Anal Appl* 428:627–652
- Xu JC (1996) Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J Numer Anal* 33:1759–1777
- Zhang T (2013) Two grid characteristic finite volume methods for nonlinear parabolic problem. *J Comput Math* 31:470–487
- Zhang GD, He YN, Zhang Y (2014) Streamline diffusion finite element method for stationary incompressible magnetohydrodynamics. *Numer Methods Partial Differ Equ* 30:1877–1901
- Zhang T, Yang JH (2014) Two level finite volume method for the unsteady Navier–Stokes equations based on two local Gauss integrations. *J Comput Appl Math* 263:377–391
- Zhang T, Yuan JY, Si ZY (2015a) Decoupled two grid finite element method for the time-dependent natural convection problem I: spatial discretization. *Numer Methods Partial Differ Equ* 31(6):2135–2168
- Zhang T, Zhao X, Huang PZ (2015b) Decoupled two level finite element methods for the steady natural convection problem. *Numer Algorithms* 68(4):837–866