Journal of Scientific Computing (2024) 98:31
https://doi.org/10.1007/s10915-023-02408-x

On Finite Volume Methods for a Navier–Stokes Variational Inequality

Feifei Jing[1] · Weimin Han[2] · Takahito Kashiwabara[3] · Wenjing Yan[4]

Received: 11 January 2023 / Revised: 11 July 2023 / Accepted: 19 August 2023

Abstract
Finite volume methods are introduced to solve a class of variational inequality of the second kind, which is governed by the incompressible Navier–Stokes equations under nonlinear slip boundary conditions of friction type. The numerical schemes are derived with the use of a proper numerical integration formula (e.g., trapezoidal rule) to approximate the subdifferential term due to the nonlinear constraints. For two particular finite volume schemes, existence and uniqueness of the numerical solutions are proved and optimal order error estimates are derived for solving the Navier–Stokes variational inequality. Numerical results are reported, focusing on the numerical convergence orders and illustrations of the tangential velocity components characterizing the potential slip phenomena.

Keywords Navier–Stokes variational inequality · Finite volume methods · Trapezoidal rule · Equivalence · Optimal order error estimate

1 Introduction

Nonlinear slip or leak boundary conditions of the friction type for the incompressible hydrodynamic equations were first introduced by Fujita [16, 17]. The boundary conditions are expressed by monotone relations between physical quantities, and the weak formulations of the problems are given in the form of variational inequalities (VIs) of second kind. Modeling, mathematical analysis and numerical approximation of such problems have been the research topic of various publications, cf. [18, 29, 30, 35, 37, 42] on Stokes variational inequalities and [14, 24, 26, 27, 36] on Navier–Stokes variational inequalities.

The first author was supported by the National Natural Science Foundation of China (No. 12001413), the Postdoctoral Science Foundation of China (No. 2021M692647) and the Fundamental Research Funds for the Central Universities of China (No. D5000230046). The second author was supported by Simons Foundation Collaboration Grants, No. 850737, and by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie Grant Agreement No. 823731 CONMECH. The third author was supported by a Grant-in-Aid for Early-Career Scientists (No. 20K14357) of the Japan Society for the Promotion of Science (JSPS). The fourth author was supported by the National Natural Science Foundation of China (No. 11971377).

✉ Wenjing Yan
wenjingyan@mail.xjtu.edu.cn

Extended author information available on the last page of the article

Published online: 03 January 2024

The problem to be investigated in this paper is described as follows. Let $\Omega \subset R^2$ be an open bounded convex polygon with a Lipschitz boundary $\partial\Omega$ consisting of two components Γ_D and Γ_S: $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_S$, $\Gamma_D \cap \Gamma_S = \emptyset$ with both Γ_D and Γ_S non-empty. We consider the two dimensional stationary incompressible flow problem:

$$-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega, \tag{1}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in } \Omega, \tag{2}$$

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_D, \tag{3}$$

$$\mathrm{u}_{\mathbf{n}} = 0, \quad |\boldsymbol{\sigma}_{\boldsymbol{\tau}}| \le g, \quad \boldsymbol{\sigma}_{\boldsymbol{\tau}}\cdot\mathbf{u}_{\boldsymbol{\tau}} + g|\mathbf{u}_{\boldsymbol{\tau}}| = 0 \quad \text{on } \Gamma_S, \tag{4}$$

where $\mathbf{u}$ is the fluid velocity, p the pressure, $\nu > 0$ the constant viscosity, and $\mathbf{f}$ a given external force density. On Γ_S, the nonlinear constraints in (4) represent a slip boundary condition of friction type. The positive function $g : \Gamma_S \to (0,\infty)$ is called the threshold slip or barrier function. The symbols $\mathbf{n}$ stands for the unit outward normal vector whereas $\boldsymbol{\tau}$ is the unit tangent vector obtained by rotating $\mathbf{n}$ by 90 degrees counter clockwise, and $\mathrm{v}_{\mathbf{n}} = \mathbf{v}\cdot\mathbf{n}$ and $\mathbf{v}_{\boldsymbol{\tau}} = \mathbf{v} - \mathrm{v}_{\mathbf{n}}\mathbf{n}$ are the normal and tangential components for a vector $\mathbf{v}$ defined on the boundary, respectively. Denote by $\boldsymbol{\sigma}_{\boldsymbol{\tau}}(\mathbf{u}) = \nu\frac{\partial\mathbf{u}_{\boldsymbol{\tau}}}{\partial\mathbf{n}}$ the tangential component of stress vector defined on Γ_S. If $g \equiv 0$, (4) reduces to the usual slip boundary condition: $\mathrm{u}_{\mathbf{n}} = 0$, $\boldsymbol{\sigma}_{\boldsymbol{\tau}} = \mathbf{0}$. From the second and third relations in (4), we see that g represents a threshold of the tangential stress such that

$$\left.\begin{array}{ll} \text{if } |\boldsymbol{\sigma}_{\boldsymbol{\tau}}| < g, & \text{then } \mathbf{u}_{\boldsymbol{\tau}} = \mathbf{0} \\ \text{if } |\boldsymbol{\sigma}_{\boldsymbol{\tau}}| = g, & \text{then } \mathbf{u}_{\boldsymbol{\tau}} \ne \mathbf{0}, \text{ and } \boldsymbol{\sigma}_{\boldsymbol{\tau}} = -g\frac{\mathbf{u}_{\boldsymbol{\tau}}}{|\mathbf{u}_{\boldsymbol{\tau}}|} \end{array}\right\} \text{ on } \Gamma_S.$$

As a numerical method lying between the finite difference method (FDM) and the finite element method (FEM), the finite volume method (FVM) inherits some physical conservation laws locally ([7, 10, 11, 34, 38, 47]) and has been successfully developed and applied to solve incompressible fluid flow problems ([8, 9, 13, 33, 40, 41, 46, 48]). Due to the low regularity of the solution and the inequality form of the VI problems, low-order elements are preferred in their numerical solution. In [29], the finite element method is studied for solving a similar Stokes variational inequality with the treatment of the nonlinear constraint on the slip boundary by the trapezoidal rule. In comparison with the FEMs, the use of the trapezoidal rule arises more naturally as a result of the pointwise enforcement of the nonlinear constraint in the FVMs; cf. [25], where three lowest-order FVMs are considered for the Stokes VI problem.

In this work, we will employ FVMs to solve the Navier–Stokes VI corresponding to the boundary value problem (1)–(4). When the nonlinear friction boundary condition is involved, FV schemes cannot be derived by simply multiplying the differential equations by piecewise-constant test functions and integrating. In fact, we do not expect to get useful information from multiplying (4) by some test functions. One natural way to treat such nonlinear constraints in the context of the FVM is to impose them pointwisely at the center points associated with each control volume. For example, we propose to discretize the second and third relations in (4) as (the symbols are defined in Sect. 3.1):

$$|\boldsymbol{\sigma}_{h\boldsymbol{\tau}}(P)| \le g(P), \qquad \boldsymbol{\sigma}_{h\boldsymbol{\tau}}(P)\cdot\mathbf{u}_{h\boldsymbol{\tau}}(P) + g(P)|\mathbf{u}_{h\boldsymbol{\tau}}(P)| = 0 \quad \forall P \in \mathcal{N}_{hS}^{\partial},$$

with $\boldsymbol{\sigma}_{h\boldsymbol{\tau}} \in \boldsymbol{\Lambda}_h$. Introducing a normalized quantity $\boldsymbol{\lambda}_h \in \boldsymbol{\Lambda}_h$, $\boldsymbol{\lambda}_h = -\boldsymbol{\sigma}_{h\boldsymbol{\tau}}/g$ on $\mathcal{N}_{hS}^{\partial}$, we have

$$|\boldsymbol{\lambda}_h(P)| \le 1, \qquad \boldsymbol{\lambda}_h(P)\cdot\mathbf{u}_{h\boldsymbol{\tau}}(P) = |\mathbf{u}_{h\boldsymbol{\tau}}(P)| \quad \forall P \in \mathcal{N}_{hS}^{\partial}. \tag{5}$$

This is equivalent to the condition

$$\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h, \qquad \langle \boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \mathbf{u}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}_h} \le 0 \quad \forall \boldsymbol{\mu}_h \in \widetilde{\boldsymbol{\Lambda}}_h,$$

where the trapezoidal rule is exploited for $\langle \cdot, \cdot \rangle_{\boldsymbol{\Lambda}_h}$, and this representation is naturally combined into a discrete formulation VE_h. This motivates the use of the trapezoidal rule to approximate the integral on the slip boundary, which also leads to an optimal error bound of the FVM solution under some assumptions.

An outline for the rest of this work is as follows. In Sect. 2, we recall some standard notation, and give a brief summary of the variational formulation for the Navier–Stokes equations with the nonlinear slip boundary condition; a solvability result is reviewed. In Sect. 3, we introduce the notation on the nodes, edges, mesh, control volumes and function spaces, detailed derivation of two conforming finite volume schemes, as well as the equivalence between the FVM and FEM. In Sect. 4, we discuss the discrete inf-sup condition, stability, existence and uniqueness of finite volume solutions. In Sect. 5, we present optimal order error estimates for the FVM with or without the stabilizer term. In Sect. 6, we present the projection algorithm and the active/inactive set algorithm, and show the numerical results to confirm the derived convergence bound. We close with some remarks in the last section.

2 Preliminaries

For a given positive integer m, we shall use the standard Sobolev space $H^m(\Omega)$ [1]

$$H^m(\Omega) = \{\mathrm{v} \in L^2(\Omega) : \partial^k \mathrm{v} \in L^2(\Omega) \ \forall\, |k| \le m\},$$

where $k = (k_1, k_2)$, k_1 and k_2 being nonnegative integers, $|k| = k_1 + k_2$, and

$$\partial^k \mathrm{v} = \frac{\partial^{|k|} \mathrm{v}}{\partial x_1^{k_1} \partial x_2^{k_2}}.$$

It is a Hilbert space with the norm and seminorm

$$\|\mathrm{v}\|_{m,\Omega} = \Big[\sum_{0 \le |k| \le m} \int_\Omega |\partial^k \mathrm{v}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} \Big]^{1/2}, \quad |\mathrm{v}|_{m,\Omega} = \Big[\sum_{|k| = m} \int_\Omega |\partial^k \mathrm{v}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} \Big]^{1/2}.$$

For functions vanishing on the boundary $\partial\Omega$, we define

$$H_0^1(\Omega) = \{\mathrm{v} \in H^1(\Omega) : \mathrm{v}|_{\partial\Omega} = 0\}.$$

We shall also need the following space of functions with zero mean value

$$L_0^2(\Omega) = \{q \in L^2(\Omega) : \int_\Omega q = 0\}.$$

Define $\mathbf{H}^1(\Omega) = [H^1(\Omega)]^2$ and

$$\mathbf{V} = \{\mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v}|_{\Gamma_D} = \mathbf{0}, \mathrm{v}_{\mathbf{n}}|_{\Gamma_S} = 0\}, \quad \mathring{Q} = L_0^2(\Omega).$$

We equip $\mathbf{V}$ with the norm $\|\nabla \cdot\|_{0,\Omega}$. Note that $\mathbf{V} \subset \mathbf{Y} = \mathbf{Y}^* \subset \mathbf{V}^*$ with $\mathbf{Y} = [L^2(\Omega)]^2$. For simplicity, we drop Ω in the notation for norms in the rest of this paper. The scalar product and norm in $\mathring{Q}$ are the usual $L^2(\Omega)$ inner product and its norm $\|\cdot\|_0$, respectively.

For $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}$ and $q \in \mathring{Q}$, define two bilinear forms $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ and a trilinear form $c(\cdot;\cdot,\cdot)$ by the formulas

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega \nu \nabla \mathbf{u} : \nabla \mathbf{v}\, \mathrm{d}\mathbf{x}, \quad b(\mathbf{v},q) = -\int_\Omega q \nabla \cdot \mathbf{v}\, \mathrm{d}\mathbf{x}, \quad c(\mathbf{u};\mathbf{v},\mathbf{w}) = \int_\Omega (\mathbf{u}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}\, \mathrm{d}\mathbf{x}.$$

Moreover, if $\mathbf{f} \in \mathbf{Y}$, we write

$$(\mathbf{f},\mathbf{v}) = \int_\Omega \mathbf{f}\cdot\mathbf{v}\, \mathrm{d}\mathbf{x} \quad \forall \mathbf{v} \in \mathbf{V}.$$

A weak formulation of the problem (1)–(4) is the following variational inequality (cf. [17]): Find $(\mathbf{u}, p) \in \mathbf{V} \times \mathring{Q}$ such that

$$(\mathrm{VI}) \quad \begin{cases} a(\mathbf{u},\mathbf{v}-\mathbf{u}) + b(\mathbf{v}-\mathbf{u},p) + c(\mathbf{u};\mathbf{u},\mathbf{v}-\mathbf{u}) + j(\mathbf{v}_{\boldsymbol{\tau}}) - j(\mathbf{u}_{\boldsymbol{\tau}}) \geq (\mathbf{f},\mathbf{v}-\mathbf{u}) \\ \hfill \forall \mathbf{v} \in \mathbf{V}, \quad \text{(6a)} \\ b(\mathbf{u},q) = 0 \quad \forall q \in \mathring{Q}, \hfill \text{(6b)} \end{cases}$$

where

$$j(\boldsymbol{\xi}) = \int_{\Gamma_S} g\, |\boldsymbol{\xi}|\, \mathrm{d}s \quad \boldsymbol{\xi} \in \mathbf{L}^2(\Gamma_S).$$

The functional j is continuous on $\mathbf{L}^2(\Gamma_S)$ but is non-differentiable. It is known that there exists a positive constant $\beta > 0$ such that [17, 42, 45]

$$\beta \|q\|_0 \leq \sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v},q)}{\|\mathbf{v}\|_{\mathbf{V}}} \quad \forall q \in \mathring{Q}.$$

Moreover, (6) is equivalent to finding $\mathbf{u} \in \mathbf{V}^{\mathrm{div}} = \{\mathbf{u} \in \mathbf{V} : \nabla\cdot\mathbf{u} = 0 \text{ in } \Omega\}$ such that

$$(\mathrm{VI}^{\mathrm{div}}) \quad a(\mathbf{u},\mathbf{v}-\mathbf{u}) + c(\mathbf{u};\mathbf{u},\mathbf{v}-\mathbf{u}) + j(\mathbf{v}_{\boldsymbol{\tau}}) - j(\mathbf{u}_{\boldsymbol{\tau}}) \geq (\mathbf{f},\mathbf{v}-\mathbf{u}) \quad \forall \mathbf{v} \in \mathbf{V}^{\mathrm{div}}. \tag{7}$$

By the Sobolev trace theorem, we have the inequality

$$\|\mathbf{v}\|_{0,\Gamma_S} \leq \varrho_0^{-1/2} \|\mathbf{v}\|_{\mathbf{V}} \quad \forall \mathbf{v} \in \mathbf{V}, \tag{8}$$

where $\varrho_0 > 0$ is the smallest eigenvalue of the following eigenvalue problem

$$\mathbf{u} \in \mathbf{V}, \quad \int_\Omega \nabla\mathbf{u} : \nabla\mathbf{v}\, \mathrm{d}\mathbf{x} = \varrho \int_{\Gamma_S} \mathbf{u}_{\boldsymbol{\tau}} \cdot \mathbf{v}_{\boldsymbol{\tau}}\, \mathrm{d}s \quad \forall \mathbf{v} \in \mathbf{V}.$$

Existence and uniqueness of a solution to the problem (7) is guaranteed under the conditions [36]

$$\mathbf{f} \in \mathbf{Y}, \quad g \in L^2(\Gamma_S), \quad 4N(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S}) < \nu^2,$$

with

$$\|\mathbf{f}\|_{\mathbf{V}^*} = \sup_{\mathbf{v}\in\mathbf{V}} \frac{(\mathbf{f},\mathbf{v})}{\|\mathbf{v}\|_{\mathbf{V}}}, \qquad N = \sup_{\mathbf{u},\mathbf{v},\mathbf{w}\in\mathbf{V}} \frac{c(\mathbf{u};\mathbf{v},\mathbf{w})}{\|\mathbf{u}\|_{\mathbf{V}}\|\mathbf{v}\|_{\mathbf{V}}\|\mathbf{w}\|_{\mathbf{V}}}.$$

We note that

$$|(\mathbf{f},\mathbf{v}) - j(\mathbf{v}_{\boldsymbol{\tau}})| \leq (\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})\|\mathbf{v}\|_{\mathbf{V}}.$$

Moreover, the solution can be bounded as follows

$$\|\mathbf{u}\|_{\mathbf{V}} \leq \frac{2}{\nu}(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S}).$$

Introducing the Lagrangian multiplier $\boldsymbol{\lambda} := -\boldsymbol{\sigma}_{\boldsymbol{\tau}}/g$, we find that

$$\boldsymbol{\sigma}_{\boldsymbol{\tau}} = -g\boldsymbol{\lambda}, \quad |\boldsymbol{\lambda}| \le 1, \quad \boldsymbol{\lambda}\cdot\mathbf{u}_{\boldsymbol{\tau}} = |\mathbf{u}_{\boldsymbol{\tau}}| \quad \text{on } \Gamma_S. \tag{9}$$

Then the problem (6) is equivalent to the following problem [19]: Find $(\mathbf{u}, p, \boldsymbol{\lambda}) \in (\mathbf{V}, Q, \tilde{\boldsymbol{\Lambda}})$ such that

$$\text{(VE)} \begin{cases} a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + c(\mathbf{u}; \mathbf{u}, \mathbf{v}) + \langle \boldsymbol{\lambda}, \mathbf{v}_{\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}} = (\mathbf{f}, \mathbf{v}) & \forall\, \mathbf{v} \in \mathbf{V}, \\ b(\mathbf{u}, q) = 0 & \forall\, q \in \mathring{Q}, \\ \langle \mathbf{u}_{\boldsymbol{\tau}}, \boldsymbol{\mu} - \boldsymbol{\lambda} \rangle_{\boldsymbol{\Lambda}} \le 0 & \forall\, \boldsymbol{\mu} \in \widetilde{\boldsymbol{\Lambda}}, \end{cases} \tag{10}$$

where $\boldsymbol{\Lambda} = \boldsymbol{L}^2(\Gamma_S)$ is a Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{\boldsymbol{\Lambda}} = (g\cdot, \cdot)_{\boldsymbol{L}^2(\Gamma_S)}$, and

$$\tilde{\boldsymbol{\Lambda}} = \{\boldsymbol{\lambda} \in \boldsymbol{\Lambda} : |\boldsymbol{\lambda}| \le 1 \text{ a.e. on } \Gamma_S\}$$

is a closed convex subset in $\boldsymbol{\Lambda}$.

3 Finite Volume Schemes

In this section, two types of FVMs are developed to solve the problem (1)–(4). First we introduce the nodes, edges and mesh partitions.

3.1 Notation Related to Primal Meshes

For simplicity, we will assume Ω to be a polygonal domain. Let $\{\mathcal{T}_h\}$ be a locally quasi-uniform family of the domain Ω into simplices K ([12]). We denote by $|K|$ the area of K, and denote by $h_e = |e|$ the length of an edge e of K. Let $h := \max_{K\in\mathcal{T}_h} h_K$, where h_K is the diameter of the simplex K. Evidently,

$$|e| \le h_K \le h \quad \forall e \subset \partial K.$$

For sets of nodes, we introduce

- $\mathcal{N}_h = \{P\}$: the set of all the nodal vertices, $\quad \mathring{\mathcal{N}}_h$: the set of all the interior nodes,
- $\mathcal{N}_h^\partial = \mathcal{N}_h \cap \partial\Omega$: boundary nodes,
- $\mathcal{N}_{hD}^\partial = \mathcal{N}_h \cap \overline{\Gamma}_D$: boundary nodes on $\overline{\Gamma}_D$, $\quad \mathcal{N}_{hS}^\partial = \mathcal{N}_h^\partial \setminus \mathcal{N}_{hD}^\partial$: boundary nodes on $\mathring{\Gamma}_S$,
- $\mathcal{N}_{hS} = \mathring{\mathcal{N}}_h \cup \mathcal{N}_{hS}^\partial$.

For sets of edges, we define

- $\mathcal{E}_h = \{e \subset \mathbb{R}^2 : e \text{ is an edge of some } K \in \mathcal{T}_h\}$,
- $\mathcal{E}_h^\partial = \{e \in \mathcal{E}_h : e \subset \partial\Omega\}$, $\quad \mathcal{E}_{hD}^\partial = \{e \in \mathcal{E}_h : e \subset \overline{\Gamma}_D\}$, $\quad \mathcal{E}_{hS}^\partial = \{e \in \mathcal{E}_h : e \subset \overline{\Gamma}_S\}$,
- $\mathcal{E}_h^\circ = \mathcal{E}_h \setminus \mathcal{E}_h^\partial$, $\quad \mathcal{E}_{hD} = \mathcal{E}_h^\circ \cup \mathcal{E}_{hD}^\partial$, $\quad \mathcal{E}_{hS} = \mathcal{E}_h^\circ \cup \mathcal{E}_{hS}^\partial$.

Finally, we define

- ξ_P^K: barycenteric coordinate on $K \in \mathcal{T}_h$ with respect to $P \in \mathcal{N}_h(K) := \mathcal{N}_h \cap K$. Namely, $\xi_P^K \in P_1(K)$ is determined by $\xi_P^K(M) = \delta_{PM}$ (the Kronecker delta) for $M \in \mathcal{N}_h(K)$,
- $K_P = \{\mathbf{x} \in K : \xi_P^K(\mathbf{x}) \ge \xi_M^K(\mathbf{x})\ \forall M \in \mathcal{N}_h(K)\}$ for $K \in \mathcal{T}_h$ and $P \in \mathcal{N}_h(K)$.

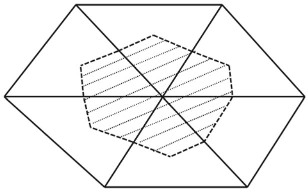


Fig. 1 A control volume at one point

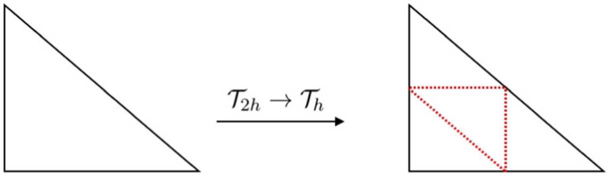


Fig. 2 An element example of the mesh generation for Case II

3.2 Function Spaces and Control Volumes

For any $P \in \mathcal{N}_h$, the control volume T_P for one vertex P are designed by connecting the barycenters and the midpoints of edges and faces of the simplex in $\mathcal{T}_h$ (Fig. 1). The cell center assigned to each control volume T_P is P. Let $\mathcal{T}_h(P) = \{K \in \mathcal{T}_h : P \in K\}$ be the collection of the triangles sharing the vertex P. Then

$$T_P = \bigcup_{K \in \mathcal{T}_h(P)} K_P.$$

The dual mesh is defined as $\mathcal{T}_h^* = \{T_P\}_{P \in \mathcal{N}_h}$. Given a non-negative integer k, $\mathcal{P}_k(K)$ denotes the space of all polynomials with a degree less than or equal to k defined on $K \in \mathcal{T}_h$. Define

$$\mathbf{V}_h = \{\mathbf{v}_h \in \mathbf{C}(\overline{\Omega}) : \mathbf{v}_h|_K \in [\mathcal{P}_1(K)]^2 \ \forall K \in \mathcal{T}_h,\ \mathbf{v}_h = \mathbf{0} \text{ on } \Gamma_D,\ \mathrm{v}_{h\mathbf{n}} = 0 \text{ on } \Gamma_S\},$$
$$\mathbf{W}_h = \{\mathbf{w}_h \in \mathbf{L}^2(\Omega) : \mathbf{w}_h|_{T_P} \in [\mathcal{P}_0(T_P)]^2 \ \forall P \in \mathcal{N}_h,\ \mathbf{w}_h|_{\mathcal{N}_{hD}^{\partial}} = \mathbf{0},\ \mathrm{w}_{h\mathbf{n}}|_{\mathcal{N}_{hS}^{\partial}} = 0\}.$$

Moreover, we introduce the sets

- $\mathbf{\Lambda}_h = \{\mathbf{w}_{h\boldsymbol{\tau}}|_{\Gamma_S} : \mathbf{w}_h \in \mathbf{W}_h\}$: the tangential traces of $\mathbf{W}_h$ to Γ_S,
- $\widetilde{\mathbf{\Lambda}}_h = \{\boldsymbol{\mu}_h \in \mathbf{\Lambda}_h : |\boldsymbol{\mu}_h| \leq 1 \text{ on } \Gamma_S\}$: a convex subset of $\mathbf{\Lambda}_h$.

For the pressure variable, two choices are listed:

$$\text{Case I:}\quad \mathring{Q}_h^0 = \{q \in \mathring{Q} : q|_K \in \mathcal{P}_0(K),\ \ \forall K \in \mathcal{T}_h\},$$
$$\text{Case II:}\quad \mathring{Q}_h^1 = \{q \in \mathring{Q} : q|_K \in \mathcal{P}_0(K),\ \ \forall K \in \mathcal{T}_{2h}\}.$$

The partition $\mathcal{T}_h$ is obtained by dividing each triangle in $\mathcal{T}_{2h}$ into four triangles by joining the midpoint of the three edges (see Fig. 2 for one element). For the lowest-order finite element pairs, $\mathbf{V}_h \times \mathring{Q}_h^0$ is unstable and a stabilizer is needed, whereas $\mathbf{V}_h \times \mathring{Q}_h^1$ is stable. We will use $\mathring{Q}_h$ to stand for $\mathring{Q}_h^0$ or $\mathring{Q}_h^1$ in the following without any ambiguity.

The transferring operator $\gamma : \mathbf{V}_h \to \mathbf{W}_h$ is given by

$$\gamma \mathbf{v}_h = \sum_{P \in \mathcal{N}_h} \mathbf{v}_h(P) \chi_{T_P},$$

where χ_{T_P} is the characteristic functions associated with the element T_P of the dual partition $\mathcal{T}_h^*$,

$$\chi_{T_P} = \begin{cases} 1 \text{ on } T_P, \\ 0 \text{ otherwise.} \end{cases}$$

3.3 Construction of the FV Schemes

To construct the FV schemes, we follow the basic strategy proposed in [25]. Multiplying the PDE by piecewise-constant test functions, integrating and performing integration by parts locally, and replacing $(\mathbf{u}, p)$ by its approximation $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h$, we obtain an equality of Petrov-Galerkin type. The next key step is to approximate the nonlinear constraints by the proper numerical integration formula, and as such, we arrive at the numerical scheme. A natural way to handle such a nonlinear constraint in the context of the FVMs is to impose the constraint pointwisely at the center points associated with each control volume, i.e., we discretize the relations along the tangential direction in (4) as

$$|\boldsymbol{\sigma}_{h\boldsymbol{\tau}}(P)| \le g(P), \quad \boldsymbol{\sigma}_{h\boldsymbol{\tau}}(P) \cdot \mathbf{u}_{h\boldsymbol{\tau}}(P) + g(P)|\mathbf{u}_{h\boldsymbol{\tau}}(P)| = 0 \quad \forall P \in \mathcal{N}_{hS}^{\partial},$$

with $\boldsymbol{\sigma}_{h\boldsymbol{\tau}} \in \boldsymbol{\Lambda}_h$. Similar to (5), with the normalized quantity $\boldsymbol{\lambda}_h = -\boldsymbol{\sigma}_{h\boldsymbol{\tau}}/g$ on $\mathcal{N}_{hS}^{\partial}$, we have

$$|\boldsymbol{\lambda}_h(P)| \le 1, \quad \boldsymbol{\lambda}_h(P) \cdot \mathbf{u}_{h\boldsymbol{\tau}}(P) = |\mathbf{u}_{h\boldsymbol{\tau}}(P)| \quad \forall P \in \mathcal{N}_{hS}^{\partial},$$

which is equivalent to the condition

$$\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h \text{ and } \langle \boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \mathbf{u}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}_h} \le 0 \quad \forall \boldsymbol{\mu}_h \in \widetilde{\boldsymbol{\Lambda}}_h,$$

where the trapezoidal rule is applied in defining $\langle \cdot, \cdot \rangle_{\boldsymbol{\Lambda}_h}$. Then this condition is incorporated in the variational formulation, cf. the saddle-point formulation VE$_h$ below. Let's first take $\mathbf{V}_h \times \mathring{Q}_h^1$ as the discrete space pair.

Step 1. For any $\mathbf{v}_h \in \mathbf{V}_h$, multiplying (1) by $\gamma \mathbf{v}_h \in \mathbf{W}_h$ and integrating over Ω, and multiplying (2) by any $q_h \in \mathring{Q}_h^1$ and integrating over Ω, we see that the exact solution satisfies

$$\begin{aligned} -\sum_{P \in \mathcal{N}_{hS}} \int_{\partial T_P \setminus \Gamma_S} \left(\nu \frac{\partial \mathbf{u}}{\partial \mathbf{n}} - p\mathbf{n} \right) \cdot \mathbf{v}_h(P)\, \mathrm{d}s + \sum_{P \in \mathcal{N}_{hS}} \int_{\partial T_P \setminus \partial\Omega} (\mathbf{u} \cdot \mathbf{n})\mathbf{u} \cdot \mathbf{v}_h(P)\, \mathrm{d}s \\ - \sum_{P \in \mathcal{N}_{hS}^{\partial}} \int_{\partial T_P \cap \Gamma_S} \boldsymbol{\sigma}_{\boldsymbol{\tau}} \cdot \mathbf{v}_{h\boldsymbol{\tau}}(P)\, \mathrm{d}s = (\mathbf{f}, \gamma \mathbf{v}) \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \\ -\int_{\Omega} \nabla \cdot \mathbf{u}\, q_h\, \mathrm{d}\mathbf{x} = 0 \quad \forall q_h \in \mathring{Q}_h^1, \\ |\boldsymbol{\sigma}_{\boldsymbol{\tau}}(P)| \le g(P), \quad \boldsymbol{\sigma}_{\boldsymbol{\tau}}(P) \cdot \mathbf{u}_{\boldsymbol{\tau}}(P) + g(P)|\mathbf{u}_{\boldsymbol{\tau}}(P)| = 0 \quad \forall P \in \mathcal{N}_{hS}^{\partial}, \end{aligned}$$

with $(\mathbf{f}, \gamma \mathbf{v}) = \sum_{T_P \in \mathcal{T}_h^*} \int_{T_P} \mathbf{f} \cdot \gamma \mathbf{v}\, \mathrm{d}\mathbf{x}$ the inner product defined on the dual mesh $\mathcal{T}_h^*$.

Step 2. Let $\boldsymbol{\Lambda}_h = \{\mathbf{w}_{h\boldsymbol{\tau}} : \mathbf{w}_h \in \mathbf{V}_h\}$. We replace $(\mathbf{u}, p, \boldsymbol{\sigma}_{\boldsymbol{\tau}})$ in the three equalities in Step 1 by $(\mathbf{u}_h, p_h, \boldsymbol{\sigma}_{h\boldsymbol{\tau}}) \in \mathbf{V}_h \times \mathring{Q}_h^1 \times \boldsymbol{\Lambda}_h$ to obtain

$$\begin{aligned} -\sum_{P \in \mathcal{N}_{hS}} \int_{\partial T_P \setminus \Gamma_S} \left(\nu \frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} - p_h\mathbf{n} \right) \cdot \mathbf{v}_h(P)\, \mathrm{d}s + \sum_{P \in \mathcal{N}_{hS}} \int_{\partial T_P \setminus \partial\Omega} (\mathbf{u}_h \cdot \mathbf{n})\mathbf{u}_h \cdot \mathbf{v}_h(P)\, \mathrm{d}s \\ - \sum_{P \in \mathcal{N}_{hS}^{\partial}} \int_{\partial T_P \cap \Gamma_S} \boldsymbol{\sigma}_{h\boldsymbol{\tau}} \cdot \mathbf{v}_{h\boldsymbol{\tau}}(P)\, \mathrm{d}s = (\mathbf{f}, \gamma \mathbf{v}_h) \quad \forall \mathbf{w}_h \in \mathbf{V}_h, \end{aligned}$$

$$-\int_\Omega \nabla\cdot\mathbf{u}_h\, q_h\,\mathrm{d}\mathbf{x}=0\quad\forall q_h\in\mathring{Q}_h^1,$$
$$|\boldsymbol{\sigma}_{h\boldsymbol{\tau}}(P)|\le g(P),\quad \boldsymbol{\sigma}_{h\boldsymbol{\tau}}(P)\cdot\mathbf{u}_{h\boldsymbol{\tau}}(P)+g(P)|\mathbf{u}_{h\boldsymbol{\tau}}(P)|=0\quad\forall\, P\in\mathcal{N}_{hS}^{\partial}.$$

Since $\boldsymbol{\sigma}_{h\boldsymbol{\tau}}$ is a piecewise constant,

$$\sum_{P\in\mathcal{N}_{hS}^{\partial}}\int_{\partial T_P\cap\Gamma_S}\boldsymbol{\sigma}_{h\boldsymbol{\tau}}\cdot\mathbf{v}_{h\boldsymbol{\tau}}(P)\,\mathrm{d}s=\sum_{e\in\mathcal{E}_{hS}^{\partial}}\frac{|e|}{2}\sum_{P\in\mathcal{N}_h\cap e}(\boldsymbol{\sigma}_{h\boldsymbol{\tau}}\cdot\mathbf{v}_{h\boldsymbol{\tau}})(P).$$

Now, using $\boldsymbol{\lambda}_h:=-\boldsymbol{\sigma}_{h\boldsymbol{\tau}}/g$ on $\mathcal{N}_{hS}^{\partial}$, we arrive at the following FV schemes:

$$-\sum_{P\in\mathcal{N}_{hS}}\int_{\partial T_P\setminus\Gamma_S}\left(\nu\frac{\partial\mathbf{u}_h}{\partial\mathbf{n}}-p_h\mathbf{n}\right)\cdot\mathbf{v}_h(P)\,\mathrm{d}s+\sum_{P\in\mathcal{N}_{hS}}\int_{\partial T_P\setminus\partial\Omega}(\mathbf{u}_h\cdot\mathbf{n})\mathbf{u}_h\cdot\mathbf{v}_h(P)\,\mathrm{d}s$$
$$+\underbrace{\sum_{e\in\mathcal{E}_{hS}^{\partial}}\frac{|e|}{2}\sum_{P\in\mathcal{N}_h\cap e}(g\boldsymbol{\lambda}_h\cdot\mathbf{v}_{h\boldsymbol{\tau}})(P)}_{=:\langle\boldsymbol{\lambda}_h,\mathbf{w}_{h\boldsymbol{\tau}}\rangle_{\boldsymbol{\Lambda}_h}}=(\mathbf{f},\gamma\mathbf{v}_h)\quad\forall\,\mathbf{v}_h\in\mathbf{V}_h,$$
$$-\int_\Omega\nabla\cdot\mathbf{u}_h\,q_h\,\mathrm{d}\mathbf{x}=0\quad\forall\, q_h\in\mathring{Q}_h^1,$$
$$|\boldsymbol{\lambda}_h(P)|\le 1,\quad\boldsymbol{\lambda}_h(P)\cdot\mathbf{u}_{h\boldsymbol{\tau}}(P)=|\mathbf{u}_{h\boldsymbol{\tau}}(P)|\quad\forall\, P\in\mathcal{N}_{hS}^{\partial}.$$

Step 3. We next rewrite the nonlinear pointwise constraint imposed at all cell centers.

Proposition 3.1 *For $\boldsymbol{\eta}_h\in\boldsymbol{\Lambda}_h$ and $\boldsymbol{\lambda}_h\in\widetilde{\boldsymbol{\Lambda}}_h$, the following relations are equivalent:*

(a) $\boldsymbol{\lambda}_h(P)\cdot\boldsymbol{\eta}_h(P)=|\boldsymbol{\eta}_h(P)|$ *for all* $P\in\mathcal{N}_{hS}^{\partial}$,
(b) $\langle\boldsymbol{\mu}_h-\boldsymbol{\lambda}_h,\boldsymbol{\eta}_h\rangle_{\boldsymbol{\Lambda}_h}\le 0\quad\forall\boldsymbol{\mu}_h\in\widetilde{\boldsymbol{\Lambda}}_h$.

Proof The implication (a) ⇒ (b) is obvious due to the inequality

$$\boldsymbol{\mu}_h(P)\cdot\boldsymbol{\eta}_h(P)\le|\boldsymbol{\mu}_h(P)|\,|\boldsymbol{\eta}_h(P)|.$$

The implication (b) ⇒ (a) follows from the fact that

$$|\boldsymbol{\eta}_h(P)|=\max\{\boldsymbol{\mu}\cdot\boldsymbol{\eta}_h(P):\boldsymbol{\mu}\in\mathbb{R}^d,\ |\boldsymbol{\mu}|\le 1\}$$

for each $P\in\mathcal{N}_{hS}^{\partial}$. □

In the FV scheme with the discrete spaces $\mathbf{V}_h\times\mathring{Q}_h^0$, we need a stabilizing term $S_h(p_h,q_h)$ which will be defined explicitly later:

$$-\int_\Omega\nabla\cdot\mathbf{u}_h\,q_h\,\mathrm{d}\mathbf{x}+S_h(p_h,q_h)=0\quad\forall q_h\in\mathring{Q}_h^0,$$

the rest of the derivation of the FV scheme remains unchanged.

We then denote

$$A(\mathbf{u}_h,\mathbf{v}_h)=-\sum_{P\in\mathcal{N}_{hS}}\int_{\partial T_P\setminus\Gamma_S}\nu\frac{\partial\mathbf{u}_h}{\partial\mathbf{n}}\cdot\mathbf{v}_h(P)\,\mathrm{d}s,$$
$$B(\mathbf{v}_h,p_h)=\sum_{P\in\mathcal{N}_{hS}}\int_{\partial T_P\setminus\Gamma_S}p_h\mathbf{n}\cdot\mathbf{v}_h(P)\,\mathrm{d}s,$$

$$\widetilde{c}(\mathbf{u}_h; \mathbf{v}_h, \mathbf{w}_h) = \sum_{P \in \mathcal{N}_{hS}} \int_{\partial T_P \setminus \partial\Omega} (\mathbf{u}_h \cdot \mathbf{n})\mathbf{v}_h \cdot \mathbf{w}_h(P)\,\mathrm{d}s,$$

$$b(\mathbf{u}_h, q_h) = -\sum_{K \in \mathcal{T}_h} \int_K \nabla \cdot \mathbf{u}_h\, q_h\,\mathrm{d}\mathbf{x}.$$

With Proposition 3.1 in mind, we obtain the following two conforming FV schemes: find $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h) \in \mathbf{V}_h \times \mathring{Q}_h \times \widetilde{\boldsymbol{\Lambda}}_h$ such that

$$(\mathrm{VE}_h^0)\begin{cases} A(\mathbf{u}_h, \mathbf{v}_h) + B(\mathbf{v}_h, p_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h) + \langle \boldsymbol{\lambda}_h, \mathbf{v}_{h\boldsymbol{\tau}}\rangle_{\Lambda_h} = (\mathbf{f}, \gamma\mathbf{v}_h) \quad \forall\, \mathbf{v}_h \in \mathbf{V}_h, & \text{(11a)} \\ b(\mathbf{u}_h, q_h) + S_h(p_h, q_h) = 0 \quad \forall\, q_h \in \mathring{Q}_h^0, & \text{(11b)} \\ \langle \boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \mathbf{u}_{h\boldsymbol{\tau}}\rangle_{\Lambda_h} \le 0 \quad \forall\, \boldsymbol{\mu}_h \in \widetilde{\Lambda}_h, & \text{(11c)} \end{cases}$$

and

$$(\mathrm{VE}_h^1)\begin{cases} A(\mathbf{u}_h, \mathbf{v}_h) + B(\mathbf{v}_h, p_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h) + \langle \boldsymbol{\lambda}_h, \mathbf{v}_{h\boldsymbol{\tau}}\rangle_{\Lambda_h} = (\mathbf{f}, \gamma\mathbf{v}_h) \quad \forall\, \mathbf{v}_h \in \mathbf{V}_h, & \text{(12a)} \\ b(\mathbf{u}_h, q_h) = 0 \quad \forall\, q_h \in \mathring{Q}_h^1, & \text{(12b)} \\ \langle \boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \mathbf{u}_{h\boldsymbol{\tau}}\rangle_{\Lambda_h} \le 0 \quad \forall\, \boldsymbol{\mu}_h \in \widetilde{\Lambda}_h. & \text{(12c)} \end{cases}$$

We will write $\mathrm{VE}_h^\bullet$ to imply either one of VE_h^0 or VE_h^1.

It is clear that the true solution $(\mathbf{u}, p, \boldsymbol{\lambda})$ satisfies the relations

$$\begin{cases} A(\mathbf{u}, \mathbf{v}_h) + B(\mathbf{v}_h, p) + \widetilde{c}(\mathbf{u}; \mathbf{u}, \mathbf{v}_h) + \langle \boldsymbol{\lambda}, \mathbf{v}_{h\boldsymbol{\tau}}\rangle_{\boldsymbol{\Lambda}_h} = (\mathbf{f}, \gamma\mathbf{v}_h) \quad \forall\, \mathbf{v}_h \in \mathbf{V}_h, \\ b(\mathbf{u}, q_h) = 0 \quad \forall\, q_h \in \mathring{Q}_h. \end{cases} \tag{13}$$

Furthermore,

$$\begin{aligned} -&\sum_{P \in \mathcal{N}_{hS}} \int_{\partial T_P \setminus \Gamma_S} \left(\nu\frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} - p_h\mathbf{n}\right) \cdot \mathbf{v}_h(P)\,\mathrm{d}s \\ &= -\sum_{K \in \mathcal{T}_h} \sum_{P \in \mathcal{N}_h(K)} \int_{\partial K_P \setminus \partial K} \left(\frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} - p_h\mathbf{n}\right) \cdot \mathbf{v}_h(P)\,\mathrm{d}s \\ &= \sum_{K \in \mathcal{T}_h} \sum_{P \in \mathcal{N}_h(K)} \int_{\partial K_P \cap \partial K} \left(\frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} - p_h\mathbf{n}\right) \cdot \mathbf{v}_h(P)\,\mathrm{d}s \\ &= \sum_{K \in \mathcal{T}_h} \int_{\partial K} \left(\frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} - p_h\mathbf{n}\right)\mathbf{n} \cdot \mathbf{v}_h\,\mathrm{d}s \\ &= \sum_{K \in \mathcal{T}_h} \int_K (\nabla\mathbf{u}_h + p_h\mathbb{I}) : \nabla\mathbf{v}_h\,\mathrm{d}\mathbf{x}, \end{aligned}$$

where we have used $\nabla \cdot (\nabla\mathbf{u}_h + p_h\mathbb{I}) = \mathbf{0}$ in K_P and trapezoidal formulas on $e \in \mathcal{E}_h$.

Proposition 3.2 *For* $\mathbf{u}_h, \mathbf{v}_h \in \mathbf{V}_h$ *and* $p_h \in \mathring{Q}_h$*, we have*

$$A(\mathbf{u}_h, \mathbf{v}_h) + B(\mathbf{v}_h, p_h) = \sum_{K \in \mathcal{T}_h} \int_K \left(\nu\nabla\mathbf{u}_h : \nabla\mathbf{v}_h - p_h\nabla \cdot \mathbf{v}_h\right)\mathrm{d}\mathbf{x} = a(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h).$$

This gives a similar finite element formulation of VE_h with $\mathbf{w}_h = \gamma \mathbf{v}_h$: find $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h) \in \mathbf{V}_h \times \mathring{Q}_h \times \widetilde{\boldsymbol{\Lambda}}_h$ such that for all $(\mathbf{v}_h, q_h, \boldsymbol{\mu}_h) \in \mathbf{V}_h \times \mathring{Q}_h \times \widetilde{\boldsymbol{\Lambda}}_h$, we have

$$
(\mathrm{VE}_h^{\prime\bullet}) \begin{cases} a(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h) + \langle \boldsymbol{\lambda}_h, \mathbf{v}_{h\boldsymbol{\tau}} \rangle_{\Lambda_h} = (\mathbf{f}, \gamma \mathbf{v}_h), & \text{(14a)} \\ b(\mathbf{u}_h, q_h) + \epsilon_1 S_h(p_h, q_h) = 0, & \text{(14b)} \\ \langle \boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \mathbf{u}_{h\boldsymbol{\tau}} \rangle_{\Lambda_h} \leq 0. & \text{(14c)} \end{cases}
$$

Here, $\epsilon_1 = 1$ if $\mathring{Q}_h = \mathring{Q}_h^0$, while $\epsilon_1 = 0$ if $\mathring{Q}_h = \mathring{Q}_h^1$, corresponding to Problem (VE_h^0) and Problem (VE_h^1), respectively.

Remark 3.1 (i) In the derivation of the finite volume scheme (Step 1), we can also adopt

$$
\widetilde{c}(\mathbf{u}_h; \mathbf{v}_h, \mathbf{w}_h) = \Big((\mathbf{u}_h \cdot \nabla)\mathbf{v}_h + \frac{1}{2}(\operatorname{div} \mathbf{u}_h)\mathbf{v}_h, \mathbf{w}_h \Big) \quad \forall\, \mathbf{u}_h, \mathbf{v}_h, \mathbf{w}_h \in \mathbf{V}_h
$$

as the trilinear term [31, 33, 43], correspondingly, $c(\mathbf{u}; \mathbf{v}, \mathbf{w}) = \Big((\mathbf{u} \cdot \nabla)\mathbf{v} + \frac{1}{2}(\operatorname{div} \mathbf{u})\mathbf{v}, \mathbf{w} \Big)$.

(ii) The nonlinear term can be recovered as $\widetilde{c}(\mathbf{u}_h; \mathbf{v}_h, \mathbf{w}_h) = c(\mathbf{u}_h; \mathbf{v}_h, \mathbf{w}_h)$ in the subspaces $\mathbf{V}_h \times \mathring{Q}_h^1$.

A stabilizing term is included in (11b) and (14b) due to the fact that the lowest order FE pair $(\mathbf{u}_h, p_h)$ does not satisfy the so-called discrete inf-sup condition. We select two types of the stabilizer and for their more applications, one can refer to [15, 32, 48].

Projection stabilizer ([4]). Let

$$
\Pi_h : L^2(\Omega) \to P_1 = \{ q \in C^0(\Omega) \cap \mathring{Q} : q|_K \in \mathcal{P}_1(K),\ \forall\, K \in \mathcal{T}_h \}
$$

be the projection operator, which satisfies the properties:

$$
\begin{cases} \|\Pi_h p\|_0 \leq C\, \|p\|_0 \quad \forall\, p \in \mathring{Q}, \\ \|\Pi_h p - p\|_0 \leq Ch \|p\|_1 \quad \forall\, p \in H^1(\Omega) \cap \mathring{Q}. \end{cases}
$$

Then the term $S_h(p, q)$ is defined by the formula

$$
S_h(p, q) = (p - \Pi_h p, q - \Pi_h q), \quad p, q \in \mathring{Q}.
$$

By Lemma 2.3 in [4],

$$
\|q_h - \Pi_h q_h\|_0 \geq C_0 h^{1/2} \| [q_h] \|_0. \tag{15}
$$

By the standard finite element theory ([4, 32]),

$$
\|p - \Pi_h p\|_0 \leq ch \|p\|_1, \quad \forall p \in H^1(\Omega). \tag{16}
$$

Jump stabilizer [2]. We define

$$
S_h(p, q) = \sum_{e \in \mathcal{E}_h^\circ} \int_e \eta_e h_e [p][q] \,\mathrm{d}s,
$$

where $\eta_e > 0$ is a parameter and properly chosen to ensure the stability of the numerical scheme.

Next, we will show the equivalence of $\mathrm{VE}_h^\bullet$ and its inequality form.

Proposition 3.3 *Problem* ($\mathrm{VE}_h^\bullet$) *is equivalent to: find* $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h$ *such that*

$$(\mathrm{VI}_h^\bullet)\begin{cases} A(\mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + B(\mathbf{v}_h - \mathbf{u}_h, p_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \\ \qquad\qquad \geq (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h) \quad \forall\, \mathbf{v}_h \in \mathbf{V}_h, & (17a) \\ b(\mathbf{u}_h, q_h) + \epsilon_1 S_h(p_h, q_h) = 0 \;\; \forall\, q_h \in \mathring{Q}_h, & (17b) \end{cases}$$

where $j_h(\cdot)$ *is defined as*

$$j_h(\boldsymbol{\eta}) := \sum_{e \in \mathcal{E}_{hS}^{\partial}} \frac{|e|}{2} \sum_{P \in \mathcal{N}_h \cap e} g|\boldsymbol{\eta}|(P).$$

Before proving this result, we describe relations between the inner product in $\boldsymbol{\Lambda}_h$ and the functional $j_h(\cdot)$, we refer to [29, Lemma 3.2] for the slip boundary condition under stable FE pairs and [28, Lemma 2.3] for leak boundary condition under stable and unstable FE pairs. We also quote the mesh-dependent inf-sup condition to deduce the unique existence of a Lagrangian multiplier $\boldsymbol{\lambda}_h \in \boldsymbol{\Lambda}_h$.

Lemma 3.1 ([28, 29])

(i) *Every* $\boldsymbol{\eta}_h \in \boldsymbol{\Lambda}_h$ *admits an extension* $\mathbf{u}_h \in \mathbf{V}_h$ *such that* $\mathbf{u}_{h\boldsymbol{\tau}} = \boldsymbol{\eta}_h$ *on* Γ_S.
(ii) *Let* $\boldsymbol{\eta}_h \in \boldsymbol{\Lambda}_h$ *and* $\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h$. *Then* $\langle \boldsymbol{\eta}_h, \boldsymbol{\lambda}_h \rangle_{\boldsymbol{\Lambda}_h} \leq j_h(\boldsymbol{\eta}_h)$.
(ii) *Let* $\boldsymbol{\eta}_h \in \boldsymbol{\Lambda}_h$ *and* $\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h$. *Then the following statements are equivalent:*

- *(a)* $\langle \boldsymbol{\eta}_h, \boldsymbol{\lambda}_h \rangle_{\boldsymbol{\Lambda}_h} = j_h(\boldsymbol{\eta}_h)$.
- *(b)* $\langle \boldsymbol{\eta}_h, \boldsymbol{\mu}_h - \boldsymbol{\lambda}_h \rangle_{\boldsymbol{\Lambda}_h} \leq 0 \;\forall \boldsymbol{\mu}_h \in \widetilde{\boldsymbol{\Lambda}}_h$.
- *(c)* $\boldsymbol{\eta}_h \cdot \boldsymbol{\lambda}_h = |\boldsymbol{\eta}_h|$ *at* $\mathcal{N}_{hS}^{\partial}$.
- *(d)* $\boldsymbol{\lambda}_h = \widetilde{P}_h(\boldsymbol{\lambda}_h + \rho\boldsymbol{\eta}_h)$ *for all* $\rho > 0$.

(iv) *Let* $\boldsymbol{\lambda}_h \in \boldsymbol{\Lambda}_h$. *Then* $\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h$ *if and only if* $\langle \boldsymbol{\eta}_h, \boldsymbol{\lambda}_h \rangle_{\boldsymbol{\Lambda}_h} \leq j_h(\boldsymbol{\eta}_h)$ *for all* $\boldsymbol{\eta}_h \in \boldsymbol{\Lambda}_h$.

Lemma 3.2 ([28, 29]) *There exists a positive constant* β_h *depending on h such that*

$$\beta_h \|\boldsymbol{\eta}_h\|_{\boldsymbol{\Lambda}_h} \leq \sup_{\mathbf{v}_h \in \mathbf{V}_h} \frac{\langle \mathbf{v}_{h\boldsymbol{\tau}}, \boldsymbol{\eta}_h \rangle_{\boldsymbol{\Lambda}_h}}{\|\mathbf{v}_h\|_{\mathbf{V}}} \quad \forall\, \boldsymbol{\eta}_h \in \boldsymbol{\Lambda}_h.$$

Proofs of these two lemmas and more details can be found in Lemmas 2.1–2.4 of [28] or Lemmas 3.1–3.3 of [29].

Proof of Proposition 3.3 If $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h) \in \mathbf{V}_h \times \mathring{Q}_h \times \widetilde{\boldsymbol{\Lambda}}_h$ is a solution of ($\mathrm{VE}_h^\bullet$), then

$$\begin{aligned} & A(\mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + B(\mathbf{v}_h - \mathbf{u}_h, p_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \\ &\quad - (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h) \\ &\quad = -\langle \boldsymbol{\lambda}_h, \mathbf{v}_{h\boldsymbol{\tau}} - \mathbf{u}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}_h} + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \\ &\quad = \sum_{e \in \mathcal{E}_{hS}^{\partial}} \frac{|e|}{2} \sum_{P \in \mathcal{N}_h \cap e} \Big[\underbrace{\big(g(-\boldsymbol{\lambda}_h \cdot \mathbf{v}_{h\boldsymbol{\tau}} + |\mathbf{v}_{h\boldsymbol{\tau}}|)\big)(P)}_{\geq 0} + \underbrace{\big(g(\boldsymbol{\lambda}_h \cdot \mathbf{u}_{h\boldsymbol{\tau}} - |\mathbf{u}_{h\boldsymbol{\tau}}|)\big)(P)}_{=0} \Big] \geq 0, \end{aligned}$$

which shows that $(\mathbf{u}_h, p_h)$ solves ($\mathrm{VI}_h^\bullet$).

Conversely, let $(\mathbf{u}_h, p_h)$ be a solution of ($\mathrm{VI}_h^\bullet$). Take $\mathbf{u}_h \pm \mathbf{v}_h$ as a test function in (17a), with $\mathbf{v}_h \in \mathring{\mathbf{V}}_h$ arbitrary. We have

$$A(\mathbf{u}_h, \mathbf{v}_h) + B(\mathbf{v}_h, p_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h) = (\mathbf{f}, \gamma\mathbf{v}_h) \quad \forall\, \mathbf{v}_h \in \mathring{\mathbf{V}}_h.$$

Since

$$\mathring{\mathbf{V}}_h = \{\mathbf{v}_h \in \mathbf{V}_h \mid \langle \mathbf{v}_{h\boldsymbol{\tau}}, \boldsymbol{\eta}_h \rangle_{\boldsymbol{\Lambda}_h} = 0 \, \forall \boldsymbol{\eta} \in \boldsymbol{\Lambda}_h\},$$

it follows from the inf-sup condition stated in Lemma 3.2 that there exists a unique $\boldsymbol{\lambda}_h \in \boldsymbol{\Lambda}_h$ such that (11a) holds. Plugging this into (17a) gives

$$\langle \mathbf{v}_{h\boldsymbol{\tau}} - \mathbf{u}_{h\boldsymbol{\tau}}, \boldsymbol{\lambda}_h \rangle_{\boldsymbol{\Lambda}_h} \leq j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \leq j_h(\mathbf{v}_{h\boldsymbol{\tau}} - \mathbf{u}_{h\boldsymbol{\tau}}) \quad \forall \mathbf{v}_h \in \mathbf{V}_h,$$

which means $(\boldsymbol{\eta}_h, \boldsymbol{\lambda}_h)_{\boldsymbol{\Lambda}_h} \leq j_h(\boldsymbol{\eta}_h) \; \forall \boldsymbol{\eta}_h \in \boldsymbol{\Lambda}_h$. Lemma 3.1(iv) implies that $\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h$. Taking $\mathbf{v}_h = 0, 2\mathbf{u}_h$ in (17a) and $\mathbf{v}_h = \mathbf{u}_h$ in (11a), we obtain $j_h(\mathbf{u}_{h\boldsymbol{\tau}}) = \langle \mathbf{u}_{h\boldsymbol{\tau}}, \boldsymbol{\lambda}_h \rangle_{\boldsymbol{\Lambda}_h}$. Thus (11c) (or (12c)) follows from Lemma 3.1(iii). Therefore, $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h)$ is a solution of Problem ($\mathrm{VE}_h^{\bullet}$). □

Remark 3.2 (i) In the FEM for solving (6), the choice of approximation of the functional j is somewhat flexible. In comparison, in the FVMs, the use of the trapezoidal approximation $j_h(\cdot)$ for $j(\cdot)$ is natural as a result of the pointwise enforcement of the nonlinear constraint (4).
(ii) Similarly, we can show that the FE scheme $\mathrm{VE}_h'^{\bullet}$ is equivalent to

$$(\mathrm{VI}_h'^{\bullet}) \begin{cases} a(\mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + b(\mathbf{v}_h - \mathbf{u}_h, p_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \\ \qquad \geq (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h) \quad \forall \mathbf{v}_h \in \mathbf{V}_h, & \text{(18a)} \\ b(\mathbf{u}_h, q_h) + \epsilon_1 S_h(p_h, q_h) = 0 \quad \forall q_h \in \mathring{Q}_h. & \text{(18b)} \end{cases}$$

For Problem $\mathrm{VE}_h'^{1}$ or Problem $\mathrm{VI}_h'^{1}$, we can again replace $\widetilde{c}(\mathbf{u}_h; \mathbf{v}_h, \mathbf{w}_h)$ by $c(\mathbf{u}_h; \mathbf{v}_h, \mathbf{w}_h)$.
(iii) If g is a constant, then $j_h(\mathbf{v}_{h\boldsymbol{\tau}}) = j(\gamma\mathbf{v}_{h\boldsymbol{\tau}})$ for $\mathbf{v}_h \in \mathbf{V}_h$.

4 Stability and Solvability

4.1 Discrete Inf-Sup Conditions

We first give some properties of the bilinear and trilinear forms.

Lemma 4.1 ([8, 33, 34, 47]) *For any* $\mathbf{v}, \mathbf{w} \in \mathbf{V}_h \cap \mathbf{V}$, $q \in \mathring{Q}_h$, *we have*

$$A(\mathbf{v}, \mathbf{w}) = \sum_{K \in \mathcal{T}_h} \nu(\nabla\mathbf{v}, \nabla\mathbf{w})_K, \quad B(\mathbf{v}, q) = -\sum_{K \in \mathcal{T}_h} (\nabla \cdot \mathbf{v}, q)_K.$$

Lemma 4.2 ([31, 40, 46]) *For any* $\mathbf{u}, \mathbf{v} \in \mathbf{V}$, $\mathbf{w} \in \mathbf{V}(h)$ *and* $q \in \mathring{Q}$,

$$A(\mathbf{v}, \mathbf{w}) \leq C|\mathbf{v}|_1\Big(|\mathbf{w}|_1 + (\sum_{K\in\mathcal{T}_h} h^2\|\Delta\mathbf{w}\|_{0,K}^2)^{1/2}\Big), \quad \mathbf{V}(h) := \mathbf{V}_h + \mathbf{V}\cap\mathbf{H}^2(\Omega),$$

$$B(\mathbf{v}, q) \leq C|\mathbf{v}|_1\Big(\|q\|_0 + (\sum_{K\in\mathcal{T}_h} h^2|q|_{1,K}^2)^{1/2}\Big), \quad \text{if } q|_K \in H^1(K) \text{ for } K \in \mathcal{T}_h,$$

$$\widetilde{c}(\mathbf{u}; \mathbf{v}, \mathbf{w}) \leq C\|\mathbf{u}\|_{\mathbf{V}}\|\mathbf{v}\|_{\mathbf{V}}\Big(\|\mathbf{w}\|_{\mathbf{V}} + (\sum_{K\in\mathcal{T}_h} h^2|\mathbf{w}|_{2,K}^2)^{1/2}\Big).$$

Furthermore, if $\mathbf{w} \in \mathbf{V}_h$, *then there exists a constant* N_h *such that*

$$\widetilde{c}(\mathbf{u}; \mathbf{v}, \mathbf{w}) \leq N_h\|\mathbf{u}\|_{\mathbf{V}}\|\mathbf{v}\|_{\mathbf{V}}\|\mathbf{w}\|_{\mathbf{V}}.$$

Next, define

$$\mathcal{B}_h(\mathbf{v}, p; \mathbf{w}, q) = A(\mathbf{v}, \mathbf{w}) + B(\mathbf{w}, p) + b(\mathbf{v}, q) + S_h(p, q).$$

Let us prove the inf-sup condition in $\mathbf{V}_h \times \mathring{Q}_h^0$ for this bilinear form.

Lemma 4.3 *There exists a positive constant* β *independent of* h *such that*

$$\sup_{(\mathbf{v}_h,q_h)\in\mathbf{V}_h\times\mathring{Q}_h^0} \frac{\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{\mathbf{V}} + \|q_h\|_0} \geq \beta_1(\|\mathbf{u}_h\|_{\mathbf{V}} + \|p_h\|_0) \quad \forall(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h^0.$$

Proof We first recall some results. It is known that for any $p_h \in \mathring{Q}_h^0 \subset \mathring{Q}$, there exists $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$ such that

$$\nabla\cdot\mathbf{w} = p_h, \quad \|\mathbf{w}\|_{\mathbf{V}} \leq C_1\|p_h\|_0. \tag{19}$$

Let $\mathbf{w}_h = \mathcal{L}_h\mathbf{w} \in \mathbf{V}_h$ be the first order Scott–Zhang interpolant of function $\mathbf{w}$, which satisfies

$$\|\mathbf{w}_h\|_{\mathbf{V}} \lesssim \|\mathbf{w}\|_{\mathbf{V}}, \quad \|\mathbf{w}-\mathbf{w}_h\|_0 + h\|\mathbf{w}-\mathbf{w}_h\|_{\mathbf{V}} \leq C_2h\|\mathbf{w}\|_{\mathbf{V}}. \tag{20}$$

Now for any $0 < \alpha \leq 1$, we have from Proposition 3.2, (19) and (20) that

$$\begin{aligned}
\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{u}_h - \alpha\mathbf{w}_h, p_h) &= A(\mathbf{u}_h, \gamma(\mathbf{u}_h - \alpha\mathbf{w}_h)) + B(\gamma(\mathbf{u}_h - \alpha\mathbf{w}_h), p_h)\\
&\quad + b(\mathbf{u}_h, p_h) + S_h(p_h, p_h)\\
&= \nu(\nabla\mathbf{u}_h, \nabla(\mathbf{u}_h - \alpha\mathbf{w}_h)) + \alpha(\nabla\cdot\mathbf{w}_h, p_h) + S_h(p_h, p_h)\\
&= \nu\|\mathbf{u}_h\|_{\mathbf{V}}^2 - \alpha(\nabla\mathbf{u}_h, \nabla\mathbf{w}_h) + \alpha(\nabla\cdot(\mathbf{w}_h - \mathbf{w}), p_h)\\
&\quad + \alpha\|p_h\|_0^2 + S_h(p_h, p_h)\\
&\geq (\nu - \frac{1}{2}\alpha C_1^2)\|\mathbf{u}_h\|_{\mathbf{V}}^2 + \frac{1}{2}\alpha\|p_h\|_0^2\\
&\quad + \alpha(\nabla\cdot(\mathbf{w}_h - \mathbf{w}), p_h) + S_h(p_h, p_h).
\end{aligned}$$

Next, by using the divergence formulae and noticing that $\mathbf{w}_h - \mathbf{w} \in \mathbf{H}_0^1(\Omega)$, we see that

$$\begin{aligned}
\alpha(\nabla\cdot(\mathbf{w}_h - \mathbf{w}), p_h) &= \alpha\sum_{K\in\mathcal{T}_h}\int_{\partial K}(\mathbf{w}_h - \mathbf{w})\cdot\mathbf{n}p_h\,\mathrm{d}s = \alpha\sum_{e\in\mathcal{E}_h^\circ}\int_e(\mathbf{w}_h - \mathbf{w})\cdot[p_h]\,\mathrm{d}s\\
&\leq \alpha\Big(\int_{\mathcal{E}_h^\circ} h_e^{-1}|\mathbf{w}_h - \mathbf{w}|^2\,\mathrm{d}s\Big)^{1/2}\Big(\int_{\mathcal{E}_h^\circ} h_e|[p_h]|^2\,\mathrm{d}s\Big)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&\leq \alpha C h^{-1}(\|\mathbf{w}_h-\mathbf{w}\|_0+h\|\mathbf{w}_h-\mathbf{w}\|_{\mathbf{V}})\Big(\int_{\mathcal{E}_h^\circ} h_e|[p_h]|^2\,\mathrm{d}s\Big)^{1/2}\\
&\leq \alpha C C_2\|\mathbf{w}\|_{\mathbf{V}}\Big(\int_{\mathcal{E}_h^\circ} h_e|[p_h]|^2\,\mathrm{d}s\Big)^{1/2}\leq \alpha C\|p_h\|_0\Big(\int_{\mathcal{E}_h^\circ} h_e|[p_h]|^2\,\mathrm{d}s\Big)^{1/2}\\
&\leq \frac{1}{2}\frac{(\alpha C)^2}{C_0}\|p_h\|_0^2+\frac{C_0}{2}h^{1/2}\|[p_h]\|_0, \quad \text{for Projection stabilizer}\\
\text{or }\ &\leq \frac{1}{2}(\alpha C)^2\|p_h\|_0^2+\frac{1}{2}h^{1/2}\|[p_h]\|_0, \quad \text{for Jump stabilizer}
\end{aligned}
$$

where the trace inequality, (19) and (20) are applied. With (16) or the definition of the jump stabilizer, the above inequality leads to

$$
\begin{aligned}
\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{u}_h-\alpha\mathbf{w}_h, p_h) &\geq \left(\nu-\frac{1}{2}\alpha C_1^2\right)\|\nabla\mathbf{u}_h\|_{\mathbf{V}}^2+\frac{1}{2}\alpha\left(1-\frac{\alpha C^2}{C_0}\right)\|p_h\|_0^2\\
&\quad+\frac{C_0}{2}h^{1/2}\|[p_h]\|_0\\
\text{or }\ &\geq \left(\nu-\frac{1}{2}\alpha C_1^2\right)\|\nabla\mathbf{u}_h\|_{\mathbf{V}}^2+\frac{1}{2}\alpha\left(1-\alpha C^2\right)\|p_h\|_0^2\\
&\quad+\frac{1}{2}h^{1/2}\|[p_h]\|_0.
\end{aligned}
$$

Let $\alpha_0=\min\{\frac{2\nu}{C_1^2}, \frac{C_0}{C^2}\}$ or $\alpha_0=\min\{\frac{2\nu}{C_1^2}, \frac{1}{C^2}\}$. For $\alpha<\alpha_0$, we have

$$
\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{u}_h-\alpha\mathbf{w}_h, p_h)\geq C_\alpha(\|\mathbf{u}_h\|_{\mathbf{V}}+\|p_h\|_0)^2.
$$

Noting that

$$
\|\mathbf{v}_h\|_{\mathbf{V}}+\|q_h\|_0=\|\mathbf{u}_h-\alpha\mathbf{w}_h\|_{\mathbf{V}}+\|p_h\|_0\leq C(\|\mathbf{u}_h\|_{\mathbf{V}}+\|p_h\|_0).
$$

Now, it follows from the above equation that

$$
\sup_{(\mathbf{v}_h,q_h)\in\mathbf{V}_h\times\mathring{Q}_h^0}\frac{\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{\mathbf{V}}+\|q_h\|_0}\geq\frac{\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{u}_h-\alpha\mathbf{w}_h, p_h)}{\|\mathbf{u}_h-\alpha\mathbf{w}_h\|_{\mathbf{V}}+\|p_h\|_0}\geq\beta_1(\|\mathbf{u}_h\|_{\mathbf{V}}+\|p_h\|_0).
$$

Thus, we arrive at the inf-sup condition with $\beta_1=\frac{C_0}{C}$. □

Remark 4.1 (i) The inf-sup condition in Lemma 4.3 is established in the finite dimensional spaces $\mathbf{V}_h\times\mathring{Q}_h^0$, rather than $\mathring{\mathbf{V}}_h\times\mathring{Q}_h^0$ as formerly referenced in the work of variation inequality problem, with

$$
\mathring{\mathbf{V}}_h=\{\mathbf{v}_h\in\mathbf{V}_h \mid \langle\mathbf{v}_{h\boldsymbol{\tau}}, \boldsymbol{\eta}_h\rangle_{\boldsymbol{\Lambda}_h}=0\}.
$$

For a proof of the inf-sup condition in $\mathring{\mathbf{V}}_h\times\mathring{Q}_h^0$, we refer to [2, 4, 6, 32, 48].

(ii) For the finite element scheme VE_h' in the event of showing no solicitude for the trilinear term, we also have such an inf-sup condition [4, 26, 32, 39]. We modify $\mathcal{B}_h(\cdot,\cdot;\cdot,\cdot)$ to

$$
\begin{aligned}
\mathcal{B}_h(\mathbf{v}, p; \mathbf{w}, q)&=a(\mathbf{v},\mathbf{w})-b(\mathbf{w}, p)+b(\mathbf{v}, q)+S_h(p, q),\\
&\qquad\forall(\mathbf{v}, p), (\mathbf{w}, q)\in\mathbf{V}_h\times\mathring{Q}_h^0.
\end{aligned}
$$

Then, there exists a positive constant β^* independent of h such that

$$
\sup_{(\mathbf{v}_h,q_h)\in\mathbf{V}_h\times\mathring{Q}_h^0}\frac{\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{\mathbf{V}}+\|q_h\|_0}\geq\beta^*(\|\mathbf{u}_h\|_{\mathbf{V}}+\|p_h\|_0)\quad\forall(\mathbf{u}_h, p_h)\in\mathbf{V}_h\times\mathring{Q}_h^0.
$$

For the stable finite element pair (Case II), the following inf-sup condition holds:

Lemma 4.4 ([6, 28]) *The bilinear form $b(\cdot,\cdot)$ satisfies the discrete inf-sup condition*

$$\beta_2\|q\|_0 \le \sup_{\mathbf{v}\in\mathbf{V}_h} \frac{b(\mathbf{v},q)}{\|\mathbf{v}\|_{\mathbf{V}}} \quad \forall q \in \mathring{Q}_h^1,$$

where $\beta_2 > 0$ is a positive constant independent of the mesh size h.

4.2 Existence and Uniqueness of Finite Volume Solutions

We are ready to study the existence and uniqueness of the finite volume methods.

Theorem 4.1 *If $\mathbf{f} \in \mathbf{V}^*$ and $g \in L^2(\Gamma_S)$ are given with*

$$\frac{4N_h(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu^2} < 1, \tag{21}$$

(i) Problem VI_h^0 has a unique solution $(\mathbf{u}_h, p_h) \in \mathcal{S}$, where

$$\begin{aligned}\mathcal{S} = \Big\{(\mathbf{v},q) \in \mathbf{V}_h \times \mathring{Q}_h^0 : \|\mathbf{v}\|_{\mathbf{V}} \le \frac{2(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu},\\ \|q\|_0 \le 2\beta_1^{-1}(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})\Big\}.\end{aligned}$$

(ii) If $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h^0$ is a solution of Problem VI_h^0, then there exists a unique $\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h$ such that $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h)$ solves Problem VE_h^0.

Proof (i). We rewrite Problem VI_h^0 as

$$\begin{aligned}&\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h - \mathbf{u}_h, q_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h)\\ &\quad + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \ge (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h), \forall(\mathbf{v}_h, q_h) \in \mathbf{V}_h \times \mathring{Q}_h^0.\end{aligned}$$

Let $\mathbf{w}_h \in \mathbf{V}_h$ be given. For any $(\mathbf{v}_h, q_h) \in \mathbf{V}_h \times \mathring{Q}_h^0$, we define an auxiliary problem

$$\begin{aligned}&\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h - \mathbf{u}_h, q_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}})\\ &\quad \ge (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h) - \widetilde{c}(\mathbf{w}_h; \mathbf{w}_h, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h).\end{aligned} \tag{22}$$

Since $\mathcal{B}_h(\mathbf{v}_h, q_h; \mathbf{v}_h, q_h)$ is coercive in $\mathbf{V}_h \times \mathring{Q}_h^0$, by the existence theorem of the solution to elliptic variational inequality of the second kind in finite dimensional space [21] or [35], we conclude that the discrete problem (22) admits a unique solution $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h^0$.

Taking $\mathbf{v}_h = \mathbf{0}, 2\mathbf{u}_h$ and $q_h = p_h$ in (22), we derive

$$\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{u}_h, p_h) + j_h(\mathbf{u}_{h\boldsymbol{\tau}}) = (\mathbf{f}, \gamma\mathbf{u}_h) - \widetilde{c}(\mathbf{w}_h; \mathbf{w}_h, \gamma\mathbf{u}_h). \tag{23}$$

By Lemma 4.2 and Cauchy–Schwarz inequality, we get for $(\mathbf{w}_h, p_h) \in \mathcal{S}$ that

$$\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{u}_h, p_h) \ge \nu\|\mathbf{u}_h\|_{\mathbf{V}}^2, \quad |\widetilde{c}(\mathbf{w}_h; \mathbf{w}_h, \gamma\mathbf{u}_h)| \le N_h\|\mathbf{w}_h\|_{\mathbf{V}}^2\|\mathbf{u}_h\|_{\mathbf{V}}, \tag{24}$$

$$|(\mathbf{f}, \gamma\mathbf{u}_h) - j_h(\mathbf{u}_{h\boldsymbol{\tau}})| \le (\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\ |g\|_{0,\Gamma_S})\|\mathbf{u}_h\|_{\mathbf{V}}. \tag{25}$$

Combined with (21), (23) becomes

$$
\begin{aligned}
\|\mathbf{u}_h\|_{\mathbf{V}} &\le \frac{\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S}}{\nu} + \frac{N_h\|\mathbf{w}_h\|_{\mathbf{V}}^2}{\nu} \le \frac{2(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu},\\
\|p_h\|_0 &\le \beta_1^{-1} \sup_{(\mathbf{v}_h,q_h)\in \mathbf{V}_h\times \mathring{Q}_h^0} \frac{\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{\mathbf{V}} + \|q_h\|_0}\\
&\le \beta_1^{-1}\Big((\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S}) + \frac{4N_h(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})^2}{\nu^2}\Big)\\
&\le 2\beta_1^{-1}(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S}),
\end{aligned}
$$

which imply for any $\mathbf{w}_h \in \mathcal{S}$, the solution $(\mathbf{u}_h, p_h)$ of (22) also belongs to $\mathcal{S}$.

Then we define a mapping $\mathcal{G}_h : (\mathbf{w}_h, p_h) \in \mathcal{S} \to (\mathbf{u}_h = \mathcal{G}_h\mathbf{w}_h, p_h) \in \mathcal{S}$ such that $(\mathbf{u}_h, p_h)$ is the unique solution of (22). We need to prove the mapping $\mathcal{G}_h$ is contractive.

For any $\mathbf{w}_h^1, \mathbf{w}_h^2 \in \mathcal{S}$, let $(\mathbf{u}_h^1 = \mathcal{G}_h\mathbf{w}_h^1, p_h^1)$ and $(\mathbf{u}_h^2 = \mathcal{G}_h\mathbf{w}_h^2, p_h^2)$ be the unique solutions of (22) corresponding to $\mathbf{w}_h^1, \mathbf{w}_h^2$, respectively. Then $\forall(\mathbf{v}_h, q_h) \in \mathbf{V}_h \times \mathring{Q}_h^0$,

$$
\begin{aligned}
&\mathcal{B}_h(\mathbf{u}_h^1, p_h^1; \mathbf{v}_h - \mathbf{u}_h^1, q_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}^1)\\
&\quad \ge (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h^1) - \widetilde{c}(\mathbf{w}_h^1; \mathbf{w}_h^1, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h^1),
\end{aligned}
\tag{26}
$$

and

$$
\begin{aligned}
&\mathcal{B}_h(\mathbf{u}_h^2, p_h^2; \mathbf{v}_h - \mathbf{u}_h^2, q_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}^2)\\
&\quad \ge (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h^2) - \widetilde{c}(\mathbf{w}_h^2; \mathbf{w}_h^2, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h^2),
\end{aligned}
\tag{27}
$$

Taking $\mathbf{v}_h = \mathbf{u}_h^2$ in (26) and $\mathbf{v}_h = \mathbf{u}_h^1$ in (27) and $q_h = p_h^1 - p_h^2$ and adding, we have

$$
\begin{aligned}
&\mathcal{B}_h(\mathbf{u}_h^1 - \mathbf{u}_h^2, p_h^1 - p_h^2; \mathbf{u}_h^1 - \mathbf{u}_h^2, p_h^1 - p_h^2)\\
&\quad \le \widetilde{c}(\mathbf{w}_h^2; \mathbf{w}_h^2, \gamma\mathbf{u}_h^1 - \gamma\mathbf{u}_h^2) - \widetilde{c}(\mathbf{w}_h^1; \mathbf{w}_h^1, \gamma\mathbf{u}_h^1 - \gamma\mathbf{u}_h^2)\\
&\quad = \widetilde{c}(\mathbf{w}_h^2 - \mathbf{w}_h^1; \mathbf{w}_h^2, \gamma\mathbf{u}_h^1 - \gamma\mathbf{u}_h^2) - \widetilde{c}(\mathbf{w}_h^1; \mathbf{w}_h^1 - \mathbf{w}_h^2, \gamma\mathbf{u}_h^1 - \gamma\mathbf{u}_h^2).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\nu\|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}}^2 &\le N_h\|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}}(\|\mathbf{w}_h^1\|_{\mathbf{V}} + \|\mathbf{w}_h^2\|_{\mathbf{V}})\|\mathbf{w}_h^1 - \mathbf{w}_h^2\|_{\mathbf{V}}\\
&\le \frac{4N_h(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu}\|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}}\|\mathbf{w}_h^1 - \mathbf{w}_h^2\|_{\mathbf{V}},
\end{aligned}
$$

that is,

$$
\|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}} \le \frac{4N_h(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu^2}\|\mathbf{w}_h^1 - \mathbf{w}_h^2\|_{\mathbf{V}}.
$$

Thanks to the condition (21), the operator $\mathcal{G}_h : \mathcal{S} \to \mathcal{S}$ is a contraction. Applying the Banach fixed-point theorem [3, Section 5.1], we deduce that there exists a unique $(\mathbf{u}_h, p_h) \in \mathcal{S}$ such that $(\mathcal{G}_h\mathbf{u}_h, p_h) = (\mathbf{u}_h, p_h)$ and $(\mathbf{u}_h, p_h)$ is the solution pair of Problem VI_h^0.

To show the uniqueness of the solution, let $(\mathbf{u}_h^1, p_h^1), (\mathbf{u}_h^2, p_h^2) \in \mathbf{V}_h \times \mathring{Q}_h^0$ be two solutions of Problem VI_h^0. For any $(\mathbf{v}_h, q_h) \in \mathbf{V}_h \times \mathring{Q}_h^0$,

$$
\begin{aligned}
&A(\mathbf{u}_h^1, \mathbf{v}_h - \mathbf{u}_h^1) + B(\mathbf{v}_h - \mathbf{u}_h^1, p_h^1) + \widetilde{c}(\mathbf{u}_h^1; \mathbf{u}_h^1, \mathbf{v}_h - \mathbf{u}_h^1)\\
&\quad + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}^1) \ge (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h^1),
\end{aligned}
\tag{28}
$$

$$
b(\mathbf{u}_h^1, q_h) + S_h(p_h^1, q_h) = 0,
\tag{29}
$$

and

$$A(\mathbf{u}_h^2, \mathbf{v}_h - \mathbf{u}_h^2) + B(\mathbf{v}_h - \mathbf{u}_h^2, p_h^2) + \widetilde{c}(\mathbf{u}_h^2; \mathbf{u}_h^2, \mathbf{v}_h - \mathbf{u}_h^2)$$
$$+ j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}^2) \geq (\mathbf{f}, \gamma \mathbf{v}_h - \gamma \mathbf{u}_h^2), \tag{30}$$
$$b(\mathbf{u}_h^2, q_h) + S_h(p_h^2, q_h) = 0, \tag{31}$$

Taking $\mathbf{v}_h = \mathbf{u}_h^2$ in (28) and $\mathbf{v}_h = \mathbf{u}_h^1$ in (30) and adding, taking $q_h = p_h^1 - p_h^2$ in (29) and (31) and subtracting, we have

$$A(\mathbf{u}_h^1 - \mathbf{u}_h^2, \mathbf{u}_h^1 - \mathbf{u}_h^2) + B(\mathbf{u}_h^1 - \mathbf{u}_h^2, p_h^1 - p_h^2) + \widetilde{c}(\mathbf{u}_h^1; \mathbf{u}_h^1, \mathbf{u}_h^1 - \mathbf{u}_h^2) - \widetilde{c}(\mathbf{u}_h^2; \mathbf{u}_h^2, \mathbf{u}_h^1 - \mathbf{u}_h^2)$$
$$\leq 0, \tag{32}$$
$$b(\mathbf{u}_h^1 - \mathbf{u}_h^2, p_h^1 - p_h^2) + S_h(p_h^1 - p_h^2, p_h^1 - p_h^2) = 0, \tag{33}$$

It follows from (32) that

$$A(\mathbf{u}_h^1 - \mathbf{u}_h^2, \mathbf{u}_h^1 - \mathbf{u}_h^2) \leq \widetilde{c}(\mathbf{u}_h^2; \mathbf{u}_h^1 - \mathbf{u}_h^2, \mathbf{u}_h^1 - \mathbf{u}_h^2)$$
$$- \widetilde{c}(\mathbf{u}_h^1 - \mathbf{u}_h^2; \mathbf{u}_h^2, \mathbf{u}_h^1 - \mathbf{u}_h^2) - S_h(p_h^1 - p_h^2, p_h^1 - p_h^2).$$

Thus,

$$\nu \|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}}^2 \leq N_h \|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}}^2 \|\mathbf{u}_h^1\|_{\mathbf{V}} + \|\mathbf{u}_h^2\|_{\mathbf{V}})$$
$$\leq N_h \|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}}^2 \frac{4\left(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2} \|g\|_{0,\Gamma_S}\right)}{\nu}.$$

Recalling the condition (21), we derive from the above inequality that $\|\mathbf{u}_h^1 - \mathbf{u}_h^2\|_{\mathbf{V}} = 0$, i.e., $\mathbf{u}_h^1 = \mathbf{u}_h^2$. Back to (33), there is

$$S_h(p_h^1 - p_h^2, p_h^1 - p_h^2) = 0. \tag{34}$$

For "Projection stabilizer", (34) implies that $p_h^1 - p_h^2 = \Pi_h(p_h^1 - p_h^2)$ is a piecewise constant function. Since $\Pi_h(p_h^1 - p_h^2) \in \mathring{Q}_h^1$ is continuous, we then conclude that $p_h^1 = p_h^2$. For "Jump stabilizer", (34) implies that $|[p_h^1 - p_h^2]|^2 = 0$, which leads to $p_h^1 = p_h^2$.
(ii). Let $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h^0$ is a solution of Problem VI_h^0. Taking $\mathbf{u}_h \pm \mathbf{v}_h$ with an arbitrary $\mathbf{v}_h \in \mathring{\mathbf{V}}_h$ as a test function in (17a), we get

$$\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \gamma \mathbf{u}_h) = (\mathbf{f}, \gamma \mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathring{\mathbf{V}}_h. \tag{35}$$

Therefore, since $\mathring{\mathbf{V}}_h = \{\mathbf{v}_h \in \mathbf{V}_h \mid \langle \mathbf{v}_{h\boldsymbol{\tau}}, \boldsymbol{\eta}_h \rangle_{\boldsymbol{\Lambda}_h} = 0\}$, the inf-sup condition given in Lemma 4.4 asserts the unique existence of $\boldsymbol{\lambda}_h \in \boldsymbol{\Lambda}_h$ such that

$$\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \gamma \mathbf{u}_h) + \langle \boldsymbol{\lambda}_h, \mathbf{v}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}_h} = (\mathbf{f}, \gamma \mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{V}_h. \tag{36}$$

Combining (36) with (17a), we have

$$\langle \boldsymbol{\lambda}_h, \mathbf{v}_{h\boldsymbol{\tau}} - \mathbf{u}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}_h} \leq j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \tag{37}$$

which implies that

$$\langle \boldsymbol{\lambda}_h, \mathbf{v}_{h\boldsymbol{\tau}} - \mathbf{u}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}_h} \leq j_h(\mathbf{v}_{h\boldsymbol{\tau}} - \mathbf{u}_{h\boldsymbol{\tau}}) \quad \forall \mathbf{v}_h \in \mathbf{V}_h. \tag{38}$$

Together with Lemma 3.1 (i), (38) leads to

$$\langle \boldsymbol{\lambda}_h, \boldsymbol{\eta}_h \rangle_{\boldsymbol{\Lambda}_h} \leq j_h(\boldsymbol{\eta}_h) \quad \forall \boldsymbol{\eta}_h \in \boldsymbol{\Lambda}_h.$$

Hence Lemma 3.1 (iv) implies that $\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h$, and (11a) holds. Moreover, taking $\mathbf{v}_h = 0$ in (37), we have $j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \leq \langle \boldsymbol{\lambda}_h, \mathbf{u}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}_h}$, which makes (11c) hold by Lemma 3.1 (iii). Therefore, $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h)$ is a solution of Problem VE_h^0. □

For Problem VE_h^1, $\nabla \cdot \mathbf{u}_h = 0$, we introduce a discrete divergence-free subspace $\mathbf{V}_h^{\mathrm{div}}$ of $\mathbf{V}_h$ as

$$\mathbf{V}_h^{\mathrm{div}} = \{\mathbf{v} \in \mathbf{V}_h : b(\mathbf{v}, q) = 0 \, \forall q \in \mathring{Q}_h^1\}.$$

We now modify

$$\|\mathbf{f}\|_{\mathbf{V}^*} = \sup_{\mathbf{v} \in \mathbf{V}_h^{\mathrm{div}}} \frac{(\mathbf{f}, \mathbf{v})}{\|\mathbf{v}\|_{\mathbf{V}}}.$$

We also define a subset of $\mathbf{V}_h^{\mathrm{div}}$:

$$\mathcal{S} = \left\{ \mathbf{v} \in \mathbf{V}_h^{\mathrm{div}} : \|\mathbf{v}\|_{\mathbf{V}} \leq \frac{2(\|\mathbf{f}\|_h + \varrho_0^{-1/2} \|g\|_{0,\Gamma_S})}{\nu} \right\}.$$

Theorem 4.2 *Under the assumption (21), there exists a unique solution* $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h^1$ *to* Problem VI_h^1*, and as well as a unique solution* $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h) \in \mathbf{V}_h \times \mathring{Q}_h^1 \times \widetilde{\boldsymbol{\Lambda}}_h$ *of* Problem VE_h^1.

Proof It is easy to see that a solution of Problem VI_h^1 satisfies

$$A(\mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \geq (\mathbf{f}, \gamma \mathbf{v}_h - \gamma \mathbf{u}_h) \quad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathrm{div}}. \tag{39}$$

Let $\mathcal{F}$ be a nonlinear map such that for each $\mathbf{w}_h \in \mathcal{S}$, $\widetilde{\mathbf{u}}_h := \mathcal{F}(\mathbf{w}_h)$ is given as the solution of the following linear problem:

$$A(\widetilde{\mathbf{u}}_h, \mathbf{v}_h - \widetilde{\mathbf{u}}_h) + \widetilde{c}(\mathbf{w}_h; \widetilde{\mathbf{u}}_h, \mathbf{v}_h - \widetilde{\mathbf{u}}_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\widetilde{\mathbf{u}}_h) \geq (\mathbf{f}, \gamma \mathbf{v}_h - \gamma \widetilde{\mathbf{u}}_h), \quad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathrm{div}}. \tag{40}$$

Taking $\mathbf{v}_h = \mathbf{0}, 2\widetilde{\mathbf{u}}_h$ in (40), we have

$$A(\widetilde{\mathbf{u}}_h, \widetilde{\mathbf{u}}_h) + \widetilde{c}(\mathbf{w}_h; \widetilde{\mathbf{u}}_h, \widetilde{\mathbf{u}}_h) + j_h(\widetilde{\mathbf{u}}_h) = (\mathbf{f}, \gamma \widetilde{\mathbf{u}}_h),$$

which gives

$$\nu \|\widetilde{\mathbf{u}}_h\|_{\mathbf{V}}^2 - N_h \|\mathbf{w}_h\|_{\mathbf{V}} \|\widetilde{\mathbf{u}}_h\|_{\mathbf{V}}^2 \leq \|\mathbf{f}\|_{\mathbf{V}^*} \|\widetilde{\mathbf{u}}_h\|_{\mathbf{V}} + \varrho_0^{-1/2} \|g\|_{0,\Gamma_S} \|\widetilde{\mathbf{u}}_h\|_{\mathbf{V}}.$$

Combining (21), the above inequality becomes

$$\|\widetilde{\mathbf{u}}_h\|_{\mathbf{V}} \leq \frac{\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2} \|g\|_{0,\Gamma_S}}{\nu - N_h \|\mathbf{w}_h\|_{\mathbf{V}}} \leq \frac{2(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2} \|g\|_{0,\Gamma_S})}{\nu}.$$

We can thus see $\mathcal{F}$ maps $\mathcal{S}$ to $\mathcal{S}$.

The map $\mathcal{F}$ is clearly continuous and therefore is compact in the finite dimensional space $\mathbf{V}_h^{\mathrm{div}}$. If $\kappa > 0$ and $\mathbf{w}_h$ satisfies $\widetilde{\mathbf{u}}_h = \mathcal{F}(\mathbf{w}_h) = \kappa \mathbf{w}_h$, then we have from (40) that

$$\begin{aligned} &\kappa A(\mathbf{w}_h, \mathbf{v}_h - \kappa \mathbf{w}_h) + \kappa \widetilde{c}(\mathbf{w}_h; \mathbf{w}_h, \mathbf{v}_h - \kappa \mathbf{w}_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\kappa \mathbf{w}_h) \\ &\geq (\mathbf{f}, \gamma \mathbf{v}_h - \gamma \kappa \mathbf{w}_h), \quad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathrm{div}}. \end{aligned} \tag{41}$$

Let $\mathbf{v}_h = \mathbf{0}$, $2\kappa\mathbf{w}_h$ in (41), we arrive at

$$\kappa(\nu\|\mathbf{w}_h\|_{\mathbf{V}}^2 - N_h\|\mathbf{w}_h\|_{\mathbf{V}}^3) \le \|\mathbf{f}\|_{\mathbf{V}^*}\|\mathbf{w}_h\|_{\mathbf{V}} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S}\|\mathbf{w}_h\|_{\mathbf{V}}.$$

So

$$\kappa \le \frac{2(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu\|\mathbf{w}_h\|_{\mathbf{V}}}.$$

Thus, $\kappa < 1$ holds only for any $\mathbf{w}_h$ on the boundary of the ball in $\mathbf{V}_h^{\mathrm{div}}$ centered at the origin with radius $r_{\mathbf{V}_h^{\mathrm{div}}} > 2(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})/\nu$. Consequently, the Leray–Schauder fixed point theorem implies that the nonlinear map $\mathcal{F}$ defined by (40) has a fixed point $\widetilde{\mathbf{u}}_h$:

$$\mathcal{F}(\widetilde{\mathbf{u}}_h) = \widetilde{\mathbf{u}}_h$$

in any ball centered at the origin with radius $r_{\mathbf{V}_h^{\mathrm{div}}} > 2(\|\mathbf{f}\|_{\mathbf{V}^*} + \varrho_0^{-1/2}\|g\|_{0,\Gamma_S})/\nu$. This fixed point $\widetilde{\mathbf{u}}_h$ is clearly a solution of the finite volume scheme (39), which in turn provides a solution of (12a) in $\mathbf{V}_h^{\mathrm{div}}$.

Recalling the inf-sup condition Lemma 4.4 and by the classical argument, we see that p exists. Moreover, the uniqueness of $(\mathbf{u}, p)$ can be similarly obtained as in the proof of Theorem 4.1.

Similarly, using $A(\mathbf{u}_h, \mathbf{v}_h) + B(\mathbf{v}_h, p_h)$ instead of $\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h)$ in (35), we can proceed to conclude that when $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times \mathring{Q}_h^1$ is a solution of Problem VI_h^1, there exists a unique $\boldsymbol{\lambda}_h \in \widetilde{\boldsymbol{\Lambda}}_h$ such that $(\mathbf{u}_h, p_h, \boldsymbol{\lambda}_h)$ solves Problem VE_h^1. □

5 A Priori Error Estimation

In this section, we will derive the optimal priori error estimates for the finite volume schemes Problem VE_h^0 and Problem VE_h^1, separately.

Let $\mathcal{I}_h : \mathbf{V} \cap [C^0(\Omega)]^2 \to \mathbf{V}_h$ be the usual Lagrangian interpolation operator, $\mathcal{L}_h : \mathring{Q} \to \mathring{Q}_h$ be the L^2-projection operator. These two satisfy the following properties:

Lemma 5.1 ([5, 20, 41]) *For all $\mathbf{v} \in \mathbf{H}^2(K)$, $q \in H^1(\Omega) \cap \mathring{Q}$, there hold*

$$\|\mathbf{v} - \mathcal{I}_h\mathbf{v}\|_{m,K} \le Ch_K^{2-m}\|\mathbf{v}\|_2, \;\; m = 0, 1, 2.$$
$$\|q - \mathcal{L}_h q\|_m \le Ch^{1-m}\|q\|_1, \;\; m = 0, 1.$$

Next, we recall some results and assumption which are quite necessary in the following analysis of the error bounds.

Lemma 5.2 ([28, 29]) *If $g \in C^2(\overline{\Gamma}_S)$ and $\eta_h \in \Lambda_h$ keeps its sign unchanged on every side, then*

$$|j_h(\eta_h) - j(\eta_h)| \le Ch^2\|\eta_h\|_{0,\Gamma_S}.$$

Assumption 5.1 (S1) Both $\mathbf{u}_{h\boldsymbol{\tau}}$ and $(\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}}$ have a constant sign on every side, (S2) sign$(\mathbf{u}_{\boldsymbol{\tau}})$ =sign$((\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}})$ on Γ_S, where $\mathcal{I}_h\mathbf{u}$ is the interpolator of $\mathbf{u}$.

An element $\eta_h \in \Lambda_h$ is said to keep a constant sign on every side e_i, $1 \le i \le N$, if either of the two conditions is satisfied: $\eta_h|_{e_i} \ge 0$ or $\eta_h|_{e_i} \le 0$.

5.1 Optimal Order Error Estimate for Problem VI_h^0

Theorem 5.1 *Let Assumptions 5.1 hold, and let* $(\mathbf{u}, p)$ *and* $(\mathbf{u}_h, p_h)$ *be the solutions of* Problem VI ($\Leftrightarrow$ Problem VE) *and* Problem VI_h^0 ($\Leftrightarrow$ Problem VE_h^0)*, respectively. Suppose* $(\mathbf{u}, p) \in \mathbf{H}^2(\Omega) \times H^1(\Omega)$ *and* $\mathbf{f} \in \mathbf{Y}$, $g \in C^2(\overline{\Gamma}_S)$*. Then there exists a positive constant C independent of h such that*

$$\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{V}} + \|p - p_h\|_0 \le C\,h.$$

Proof Denote $\mathcal{B}(\mathbf{u}, p; \mathbf{v}, q) = A(\mathbf{u}, \mathbf{v}) + B(\mathbf{v}, p) + b(\mathbf{u}, q)$. The true solution $(\mathbf{u}, p)$ of Problem VI satisfies

$$\mathcal{B}(\mathbf{u}, p; \mathbf{v} - \mathbf{u}, q) + \widetilde{c}(\mathbf{u}; \mathbf{u}, \mathbf{v} - \mathbf{u}) + j(\mathbf{v}_{\boldsymbol{\tau}}) - j(\mathbf{u}_{\boldsymbol{\tau}}) \ge (\mathbf{f}, \mathbf{v} - \mathbf{u}), \forall(\mathbf{v}, q) \in \mathbf{V} \times \mathring{Q}. \tag{42}$$

Problem VI_h^0 is rewritten as

$$\mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{v}_h - \mathbf{u}_h, q_h) + \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) \ge (\mathbf{f}, \gamma\mathbf{v}_h - \gamma\mathbf{u}_h). \tag{43}$$

For all $(\mathbf{v}_h, q_h) \in \mathbf{V}_h \times \mathring{Q}_h^0$, it is obvious that

$$\begin{aligned}
\nu\|\mathbf{u}_h - \mathbf{v}_h\|_{\mathbf{V}}^2 + \|p_h - q_h\|_0^2 &\le \mathcal{B}_h(\mathbf{u}_h - \mathbf{v}_h, p_h - q_h; \mathbf{u}_h - \mathbf{v}_h, p_h - q_h)\\
&= \mathcal{B}_h(\mathbf{u}_h, p_h; \mathbf{u}_h - \mathbf{v}_h, p_h - q_h)\\
&\quad - S_h(q_h, p_h - q_h) - \mathcal{B}(\mathbf{u}, p; \mathbf{u}_h - \mathbf{v}_h, p_h - q_h)\\
&\quad + \mathcal{B}(\mathbf{u} - \mathbf{v}_h, p - q_h; \mathbf{u}_h - \mathbf{v}_h, p_h - q_h).
\end{aligned} \tag{44}$$

Substituting (43) with $\mathbf{v}_h$ itself, $q_h = p_h - q_h$ and (42) with $\mathbf{v}_h = \mathbf{u}_h - \mathbf{v}_h + \mathbf{u}$, $q = p_h - q_h$ in (44), we follow an exploitation of the fact $c(\mathbf{u}; \mathbf{u}, \mathbf{v}_h) = \widetilde{c}(\mathbf{u}; \mathbf{u}, \mathbf{v}_h)$ that

$$\begin{aligned}
\nu\|\mathbf{u}_h - \mathbf{v}_h\|_{\mathbf{V}}^2 + \|p_h - q_h\|_0^2 &= \mathcal{B}(\mathbf{u} - \mathbf{v}_h, p - q_h; \mathbf{u}_h - \mathbf{v}_h, p_h - q_h) - S_h(q_h, p_h - q_h)\\
&\quad + \widetilde{c}(\mathbf{u}; \mathbf{u}, \mathbf{u}_h - \mathbf{v}_h) - \widetilde{c}(\mathbf{u}_h; \mathbf{u}_h, \mathbf{u}_h - \mathbf{v}_h)\\
&\quad + (\mathbf{f}, \gamma\mathbf{u}_h - \gamma\mathbf{v}_h) - (\mathbf{f}, \mathbf{u}_h - \mathbf{v}_h)\\
&\quad + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) + j((\mathbf{u}_h - \mathbf{v}_h + \mathbf{u})_{\boldsymbol{\tau}}) - j(\mathbf{u}_{\boldsymbol{\tau}}).
\end{aligned} \tag{45}$$

Setting $\mathbf{v}_h = \mathcal{I}_h\mathbf{u}$, $q_h = \mathcal{L}_h p$, for an arbitrarily small $\varepsilon > 0$, one sees that

$$\begin{aligned}
&\mathcal{B}(\mathbf{u} - \mathbf{v}_h, p - q_h; \mathbf{u}_h - \mathbf{v}_h, p_h - q_h)\\
&\quad \le C\Big(\|\mathbf{u} - \mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}\|\mathbf{u}_h - \mathcal{I}_h\mathbf{u}\|_{\mathbf{V}} + \|\mathbf{u}_h - \mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}\|p - \mathcal{L}_h p\|_0 + \|\mathbf{u} - \mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}\|p_h - q_h\|_0\Big)\\
&\quad \le Ch^2 + \varepsilon_1\|\mathbf{u}_h - \mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}^2 + \varepsilon_2\|p_h - \mathcal{L}_h p\|_0^2.
\end{aligned} \tag{46}$$

By the definition and properties of the stabilized term $S_h(\cdot, \cdot)$ and the operator γ, one has

$$\begin{aligned}
S_h(q_h, p_h - q_h) &= S_h(p - q_h, p_h - q_h) - S_h(p, p_h - q_h)\\
&\le \|p - \mathcal{L}_h p\|_0\|p_h - \mathcal{L}_h p\|_0 + \|p - \Pi_h p\|_0\|p_h - \mathcal{L}_h p\|_0\\
&\le Ch^2 + \varepsilon_2\|p_h - \mathcal{L}_h p\|_0^2.
\end{aligned} \tag{47}$$

$$\begin{aligned}
(\mathbf{f}, \gamma\mathbf{u}_h - \gamma\mathbf{v}_h) - (\mathbf{f}, \mathbf{u}_h - \mathbf{v}_h) &\le Ch\|\mathbf{f}\|_0\|\mathbf{u}_h - \mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}\\
&\le Ch^2 + \varepsilon_1\|\mathbf{u}_h - \mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}^2.
\end{aligned} \tag{48}$$

Using Lemmas 4.2, 5.1 and Theorem 4.1, we get

$$\begin{aligned}
&\widetilde{c}(\mathbf{u};\mathbf{u},\mathbf{u}_h-\mathbf{v}_h)-\widetilde{c}(\mathbf{u}_h;\mathbf{u}_h,\mathbf{u}_h-\mathbf{v}_h)\\
&\quad=\widetilde{c}(\mathbf{u};\mathbf{u}-\mathbf{v}_h,\mathbf{u}_h-\mathbf{v}_h)-\widetilde{c}(\mathbf{u};\mathbf{u}_h-\mathbf{v}_h,\mathbf{u}_h-\mathbf{v}_h)\\
&\qquad+\widetilde{c}(\mathbf{u}-\mathbf{v}_h;\mathbf{u}_h,\mathbf{u}_h-\mathbf{v}_h)-\widetilde{c}(\mathbf{u}_h-\mathbf{v}_h;\mathbf{u}_h,\mathbf{u}_h-\mathbf{v}_h)\\
&\quad\le N_h(\|\mathbf{u}\|_{\mathbf{V}}+\|\mathbf{u}_h\|_{\mathbf{V}})\|\mathbf{u}-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}\\
&\qquad-\widetilde{c}(\mathbf{u};\mathbf{u}_h-\mathcal{I}_h\mathbf{u},\mathbf{u}_h-\mathcal{I}_h\mathbf{u})-\widetilde{c}(\mathbf{u}_h-\mathcal{I}_h\mathbf{u};\mathbf{u}_h,\mathbf{u}_h-\mathcal{I}_h\mathbf{u})\\
&\quad\le Ch^2+\varepsilon_1\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}^2-\widetilde{c}(\mathbf{u};\mathbf{u}_h-\mathcal{I}_h\mathbf{u},\mathbf{u}_h-\mathcal{I}_h\mathbf{u})-\widetilde{c}(\mathbf{u}_h-\mathcal{I}_h\mathbf{u};\mathbf{u}_h,\mathbf{u}_h-\mathcal{I}_h\mathbf{u}).
\end{aligned}\tag{49}$$

By the definition of the nonlinear term $j(\cdot)$ and Lemma 5.2, we have

$$\begin{aligned}
&j_h(\mathbf{v}_{h\boldsymbol{\tau}})-j_h(\mathbf{u}_{h\boldsymbol{\tau}})+j((\mathbf{u}_h-\mathbf{v}_h+\mathbf{u})_{\boldsymbol{\tau}})-j(\mathbf{u}_{\boldsymbol{\tau}})\\
&\quad\le j_h(\mathbf{v}_{h\boldsymbol{\tau}})-j_h(\mathbf{u}_{h\boldsymbol{\tau}})+j(\mathbf{u}_{h\boldsymbol{\tau}})-j(\mathbf{v}_{h\boldsymbol{\tau}})\\
&\quad\le |j(\mathbf{u}_{h\boldsymbol{\tau}})-j_h(\mathbf{u}_{h\boldsymbol{\tau}})|+|j((\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}})-j_h((\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}}))|\\
&\quad\le Ch^2\|\mathbf{u}_{h\boldsymbol{\tau}}\|_{0,\Gamma_S}+Ch^2\|(\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}})\|_{0,\Gamma_S}\\
&\quad\le Ch^2\|\mathbf{u}_{h\boldsymbol{\tau}}\|_{\mathbf{V}}+Ch^2\|(\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}})\|_{\mathbf{V}}\\
&\quad\le Ch^2.
\end{aligned}\tag{50}$$

Substituting (46)–(50) in (45), one has

$$\begin{aligned}
&\nu\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}^2+\|p_h-\mathcal{L}_hp\|_0^2+\widetilde{c}(\mathbf{u};\mathbf{u}_h-\mathcal{I}_h\mathbf{u},\mathbf{u}_h-\mathcal{I}_h\mathbf{u})+\widetilde{c}(\mathbf{u}_h-\mathcal{I}_h\mathbf{u};\mathbf{u}_h,\mathbf{u}_h-\mathcal{I}_h\mathbf{u})\\
&\quad\le Ch^2+3\varepsilon_1\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}^2+2\varepsilon_2\|p_h-\mathcal{L}_hp\|_0^2.
\end{aligned}$$

Thus, it follows from (21) and the above estimate that

$$\begin{aligned}
&\nu\Big(1-\frac{4N_h(\|\mathbf{f}\|_{\mathbf{V}^*}+\varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu^2}\Big)\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}^2+\|p_h-\mathcal{L}_hp\|_0^2\\
&\quad\le Ch^2+3\varepsilon_1\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}^2+2\varepsilon_2\|p_h-\mathcal{L}_hp\|_0^2.
\end{aligned}$$

Denote $\vartheta=1-4N_h(\|\mathbf{f}\|_{\mathbf{V}^*}+\varrho_0^{-1/2}\|g\|_{0,\Gamma_S})/\nu^2$. Taking $\varepsilon_1=\nu\vartheta/6$ and $\varepsilon_2=1/4$ in the above inequality, we have

$$\nu\vartheta\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}+\|p_h-\mathcal{L}_hp\|_0\le Ch.$$

Using the triangle inequality

$$\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}+\|p-p_h\|_0\le\|\mathbf{u}-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}+\|p-\mathcal{L}_hp\|_0+\|\mathbf{u}_h-\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}+\|p_h-\mathcal{L}_hp\|_0,$$

we complete the proof of the optimal error bound for Problem VI_h^0 ($\Leftrightarrow$ Problem VE_h^0). □

5.2 Optimal Order Error Estimate for Problem VI_h^1

Theorem 5.2 *Let* $(\mathbf{u},p)\in\mathbf{V}\times\mathring{Q}$ *and* $(\mathbf{u}_h,p_h)\in\mathbf{V}_h\times\mathring{Q}_h^1$ *be the solutions of* Problem VI ($\Leftrightarrow$ ProblemVE) *and* Problem VI_h^1 ($\Leftrightarrow$ ProblemVE$_h^1$), *respectively. Assume* $(\mathbf{u},p)\in\mathbf{H}^2(\Omega)\times H^1(\Omega)$, $\mathbf{f}\in\mathbf{Y}$, $g\in C^2(\overline{\Gamma}_S)$, *and Assumption 5.1. Then for a constant* C *independent of* h,

$$\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}+\|p-p_h\|_0\le C\,h.$$

Proof For arbitrary $\mathbf{v}_h$, we begin with the following equality

$$a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h) = a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h) - a(\mathbf{u},\mathbf{u}_h-\mathbf{u}) - a(\mathbf{u}_h,\mathbf{v}_h-\mathbf{u}_h) + a(\mathbf{u},\mathbf{v}_h-\mathbf{u}).$$

Substituting (6a) with $\mathbf{v}=\mathbf{u}_h$, (18a) with $\mathbf{v}_h$ itself and (10) with $\mathbf{v}=\mathbf{v}_h-\mathbf{u}$ into the second, third and fourth terms of the right-hand side in the above equation, respectively. Combining this with Korn's inequality, we deduce that

$$\begin{aligned}
\nu\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2 \le\ & a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h)\\
&+ c(\mathbf{u};\mathbf{u},\mathbf{u}_h-\mathbf{u}) + b(\mathbf{u}_h-\mathbf{u},p) + j(\mathbf{u}_{h\boldsymbol{\tau}}) - j(\mathbf{u}_{\boldsymbol{\tau}}) - (\mathbf{f},\mathbf{u}_h-\mathbf{u})\\
&+\widetilde{c}(\mathbf{u}_h;\mathbf{u}_h,\mathbf{v}_h-\mathbf{u}_h)+b(\mathbf{v}_h-\mathbf{u}_h,p_h)+j_h(\mathbf{v}_{h\boldsymbol{\tau}})-j_h(\mathbf{u}_{h\boldsymbol{\tau}})-(\mathbf{f},\gamma\mathbf{v}_h-\gamma\mathbf{u}_h)\\
&- c(\mathbf{u};\mathbf{u},\mathbf{v}_h-\mathbf{u}) - b(\mathbf{v}_h-\mathbf{u},p) - \langle\boldsymbol{\lambda},\mathbf{v}_{h\boldsymbol{\tau}}-\mathbf{u}_{\boldsymbol{\tau}}\rangle_{\boldsymbol{\Lambda}} + (\mathbf{f},\mathbf{v}_h-\mathbf{u})\\
\le\ & a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h) + \widetilde{c}(\mathbf{u};\mathbf{u},\mathbf{u}_h-\mathbf{v}_h) + \widetilde{c}(\mathbf{u}_h;\mathbf{u}_h,\mathbf{v}_h-\mathbf{u}_h)\\
&+ b(\mathbf{u}_h-\mathbf{u},p-q_h) + b(\mathbf{v}_h-\mathbf{u},p_h-p)\\
&+ j(\mathbf{u}_{h\boldsymbol{\tau}}) - j(\mathbf{u}_{\boldsymbol{\tau}}) + j_h(\mathbf{v}_{h\boldsymbol{\tau}}) - j_h(\mathbf{u}_{h\boldsymbol{\tau}}) - \langle\boldsymbol{\lambda},\mathbf{v}_{h\boldsymbol{\tau}}-\mathbf{u}_{\boldsymbol{\tau}}\rangle_{\boldsymbol{\Lambda}}.
\end{aligned}\tag{51}$$

Take $\mathbf{u}\pm\mathbf{v}_h$ as a test function in (6a), with an arbitrary $\mathbf{v}_h\in\overset{\circ}{\mathbf{V}}_h$, to get

$$a(\mathbf{u},\mathbf{v}_h) + c(\mathbf{u};\mathbf{u},\mathbf{v}_h) + b(\mathbf{v}_h,p) = (\mathbf{f},\mathbf{v}_h).$$

On the other hand, we can similarly show that

$$a(\mathbf{u}_h,\mathbf{v}_h) + \widetilde{c}(\mathbf{u}_h;\mathbf{u}_h,\mathbf{v}_h) + b(\mathbf{v}_h,p_h) = (\mathbf{f},\gamma\mathbf{v}_h) \quad \forall\,\mathbf{v}_h\in\overset{\circ}{\mathbf{V}}_h.$$

By subtraction we obtain

$$a(\mathbf{u}-\mathbf{u}_h,\mathbf{v}_h) + \widetilde{c}(\mathbf{u};\mathbf{u},\mathbf{v}_h) - \widetilde{c}(\mathbf{u}_h;\mathbf{u}_h,\mathbf{v}_h) + b(\mathbf{v}_h,p-p_h) = (\mathbf{f},\mathbf{v}_h-\gamma\mathbf{v}_h) \quad \forall\,\mathbf{v}_h\in\overset{\circ}{\mathbf{V}}_h. \tag{52}$$

It follows from Lemma 4.4, (48) and (49) that

$$\begin{aligned}
&\|p_h-q_h\|_0\\
&\le \frac{1}{\beta_2}\sup_{\mathbf{v}_h\in\overset{\circ}{\mathbf{V}}_h}\frac{b(\mathbf{v}_h,p_hq_h)}{\|\mathbf{v}\|_{\mathbf{V}}} \le \frac{1}{\beta_2}\sup_{\mathbf{v}_h\in\overset{\circ}{\mathbf{V}}_h}\frac{b(\mathbf{v}_h,p_h-p)}{\|\mathbf{v}_h\|_{\mathbf{V}}}\\
&\le \frac{1}{\beta_2}\sup_{\mathbf{v}_h\in\overset{\circ}{\mathbf{V}}_h}\frac{b(\mathbf{v}_h,p_h-p)+b(\mathbf{v}_h,p-q_h)}{\|\mathbf{v}_h\|_{\mathbf{V}}}\\
&\le \frac{1}{\beta_2}\sup_{\mathbf{v}_h\in\overset{\circ}{\mathbf{V}}_h}\frac{b(\mathbf{v}_h,pq_h)+a(\mathbf{u}-\mathbf{u}_h,\mathbf{v}_h)+\widetilde{c}(\mathbf{u};\mathbf{u},\mathbf{v}_h)-\widetilde{c}(\mathbf{u}_h;\mathbf{u}_h,\mathbf{v}_h)-(\mathbf{f},\mathbf{v}_h-\gamma\mathbf{v}_h)}{\|\mathbf{v}_h\|_{\mathbf{V}}}\\
&\le C\Big(\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}} + \|p-q_h\|_0 + h\Big).
\end{aligned}$$

Now it is clear that

$$\|p-p_h\|_0 \le \|p-q_h\|_0 + \|p_h-q_h\|_0 \le C\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}} + C\|p-q_h\|_0 + Ch. \tag{53}$$

Recall the interpolation error estimates Lemma 5.1. Taking $\mathbf{v}_h=\mathcal{I}_h\mathbf{u}$, $q_h=\mathcal{L}_hp$ in (51) and using the fact that $c(\mathbf{u};\mathbf{u},\mathbf{v}_h)=\widetilde{c}(\mathbf{u};\mathbf{u},\mathbf{v}_h)$, we have

$$|a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathcal{I}_h\mathbf{u})| \le Ch\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}} \le Ch^2 + \varepsilon_3\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2. \tag{54}$$

$$|b(\mathbf{u}_h-\mathbf{u},p-\mathcal{L}_hp)| \le Ch\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}} \le Ch^2 + \varepsilon_3\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2. \tag{55}$$

$$|b(\mathcal{I}_h\mathbf{u}-\mathbf{u},\, p_h-p)| \leq Ch\|p-p_h\|_0 \leq Ch^2+\varepsilon_3\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}. \tag{56}$$

$$\begin{aligned}
&\widetilde{c}(\mathbf{u};\mathbf{u},\mathbf{u}_h-\mathbf{v}_h)+\widetilde{c}(\mathbf{u}_h;\mathbf{u}_h,\mathbf{v}_h-\mathbf{u}_h)=\widetilde{c}(\mathbf{u};\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h-\mathbf{v}_h)+\widetilde{c}(\mathbf{u}-\mathbf{u}_h;\mathbf{u}_h,\mathbf{u}_h-\mathbf{v}_h)\\
&\quad=\widetilde{c}(\mathbf{u};\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h)+\widetilde{c}(\mathbf{u}-\mathbf{u}_h;\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h)\\
&\qquad-\widetilde{c}(\mathbf{u};\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h)-\widetilde{c}(\mathbf{u}-\mathbf{u}_h;\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h)\\
&\quad\leq Ch^2+\varepsilon_3\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2-\widetilde{c}(\mathbf{u};\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h)-\widetilde{c}(\mathbf{u}-\mathbf{u}_h;\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h).
\end{aligned}\tag{57}$$

Since $\boldsymbol{\sigma}_{\boldsymbol{\tau}}\cdot\mathbf{u}_{\boldsymbol{\tau}}+g|\mathbf{u}_{\boldsymbol{\tau}}|=0$ and $\boldsymbol{\sigma}_{\boldsymbol{\tau}}=-g\boldsymbol{\lambda}$, from [29] we know that

$$\langle\boldsymbol{\lambda},(\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}}-\mathbf{u}_{\boldsymbol{\tau}}\rangle_{\boldsymbol{\Lambda}}+j((\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}})-j(\mathbf{u}_{\boldsymbol{\tau}})=0.$$

Therefore,

$$|j(\mathbf{u}_{h\boldsymbol{\tau}})-j_h(\mathbf{u}_{h\boldsymbol{\tau}})|\leq Ch^2\|\mathbf{u}_{h\boldsymbol{\tau}}\|_{0,\Gamma_S}\leq Ch^2\|\mathbf{u}_h\|_{\mathbf{V}}\leq Ch^2. \tag{58}$$

$$|j_h((\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}})-j((\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}})|\leq Ch^2\|(\mathcal{I}_h\mathbf{u})_{\boldsymbol{\tau}}\|_{0,\Gamma_S}\leq Ch^2\|\mathcal{I}_h\mathbf{u}\|_{\mathbf{V}}\leq Ch^2. \tag{59}$$

Substituting (54)–(59) in (51),

$$\nu\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2+\widetilde{c}(\mathbf{u};\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h)+\widetilde{c}(\mathbf{u}-\mathbf{u}_h;\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h)\leq Ch^2+4\varepsilon_3\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2.$$

Thus, from (21) and the above estimate turns into

$$\nu\Big(1-\frac{4N_h(\|\mathbf{f}\|_{\mathbf{V}^*}+\varrho_0^{-1/2}\|g\|_{0,\Gamma_S})}{\nu^2}\Big)\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2\leq Ch^2+4\varepsilon_3\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}^2.$$

Again, let $\vartheta=1-4N_h(\|\mathbf{f}\|_{\mathbf{V}^*}+\varrho_0^{-1/2}\|g\|_{0,\Gamma_S})/\nu^2$ and $\varepsilon_3=\nu\vartheta/8$, we arrive at

$$\nu\|\mathbf{u}-\mathbf{u}_h\|_{\mathbf{V}}\leq Ch.$$

Now from (53), we know $\|p-p_h\|_0\leq Ch$, and the optimal error bound is arrived. □

6 Numerical Results

6.1 Numerical Algorithms

We will present two approaches to solve Problem VE_h^0 and Problem VE_h^1. The first one is based on the classical Uzawa method, which is also called the projection method ([22, 29, 37, 49]).

Algorithm 1 The projection algorithm

Step 1. Choose an arbitrary $\boldsymbol{\lambda}_h^0\in\boldsymbol{\Lambda}_h$ and $\rho>0$.

Step 2. With $\boldsymbol{\lambda}_h^n$ known, determine $(\mathbf{u}_h^n,p_h^n)\in\mathbf{V}_h\times\mathring{Q}_h$ such that

$$\begin{cases} a(\mathbf{u}_h^n,\mathbf{v}_h)+\widetilde{c}(\mathbf{u}_h^n;\mathbf{u}_h^n,\mathbf{v}_h)+b(\mathbf{v}_h,p_h^n)=(\mathbf{f},\gamma\mathbf{v}_h)-\langle\boldsymbol{\lambda}_h^n,\mathbf{v}_{h\boldsymbol{\tau}}\rangle_{\boldsymbol{\Lambda}} & \forall\,\mathbf{v}_h\in\mathbf{V}_h,\\ b(\mathbf{u}_h^n,q_h)+\epsilon_1 S_h(p_h^n,q_h)=0 \quad \forall\,q_h\in\mathring{Q}_h. \end{cases}$$

Step 3. Set $\boldsymbol{\lambda}_h^{n+1}=P_{\boldsymbol{\Lambda}}(\boldsymbol{\lambda}_h^n+\rho g\mathbf{u}_{h\boldsymbol{\tau}}^n)$, with $P_{\boldsymbol{\Lambda}}(\boldsymbol{\mu})=\sup\{-\mathbf{1},\inf\{\mathbf{1},\boldsymbol{\mu}\}\}$.

Step 4. Go to Step 2 and 3 until $\|\mathbf{u}_h^{(n)}-\mathbf{u}_h^{(n-1)}\|_V<\varepsilon_{\mathrm{tol}}$.

The second algorithm is motivated by a primal-dual active set strategy known in optimal control theory [28, 44].

Algorithm 2 The active/inactive set algorithm

Step 1. Choose $\rho > 0$ and set $\boldsymbol{\lambda}_h^0 \in \boldsymbol{\Lambda}_h$. find $(\mathbf{u}_h^0, p_h^0) \in \mathbf{V}_h \times \mathring{Q}_h$ such that

$$\begin{cases} a(\mathbf{u}_h^0, \mathbf{v}_h) + \widetilde{c}(\mathbf{u}_h^0; \mathbf{u}_h^0, \mathbf{v}_h) + b(\mathbf{v}_h, p_h^0) = (\mathbf{f}, \gamma \mathbf{v}_h) - \langle \boldsymbol{\lambda}_h^0, \mathbf{v}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}} \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \\ b(\mathbf{u}_h^0, q_h) + \epsilon_1 S_h(p_h^0, q_h) = 0 \quad \forall q_h \in \mathring{Q}_h. \end{cases}$$

Step 2. For $n \geq 1$, with $(\mathbf{u}_h^{n-1}, p_h^{n-1}, \boldsymbol{\lambda}_h^{n-1})$ known, define the active set A_n by

$$A_n = \{M \in \mathcal{N}_{hS}^{\partial} \mid |\boldsymbol{\lambda}_h^{n-1} + \rho g \mathbf{u}_h^{n-1}| \geq 1 \ at \ M\}.$$

Determine $\boldsymbol{\lambda}_h^n$ on A_n by

$$\boldsymbol{\lambda}_h^n = P_{\boldsymbol{\Lambda}}(\boldsymbol{\lambda}_h^{n-1} + \rho g \mathbf{u}_{h\boldsymbol{\tau}}^{n-1}) \text{ on } A_n.$$

Step 3. Define the inactive set $I_n = \mathcal{N}_{hS}^{\partial} \setminus A_n$. Determine $(\mathbf{u}_h^n, p_h^n)$ such that $\mathbf{u}_{h\boldsymbol{\tau}} = \mathbf{0}$ on I_n and

$$\begin{cases} a(\mathbf{u}_h^n, \mathbf{v}_h) + \widetilde{c}(\mathbf{u}_h^n; \mathbf{u}_h^n, \mathbf{v}_h) + b(\mathbf{v}_h, p_h^n) = (\mathbf{f}, \gamma \mathbf{v}_h) - \langle \boldsymbol{\lambda}_h^n, \mathbf{v}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}} \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \ \mathbf{v}_{h\boldsymbol{\tau}} = \mathbf{0} \text{ on } I_n, \\ b(\mathbf{u}_h^n, q_h) + \epsilon_1 S_h(p_h^n, q_h) = 0 \quad \forall q_h \in \mathring{Q}_h. \end{cases}$$

Determine $\boldsymbol{\lambda}_h^n$ at I_k in the following way

$$\langle \boldsymbol{\lambda}_h^n, \mathbf{v}_{h\boldsymbol{\tau}} \rangle_{\boldsymbol{\Lambda}} = (\mathbf{f}, \gamma \mathbf{v}_h) - a(\mathbf{u}_h^n, \mathbf{v}_h) - \widetilde{c}(\mathbf{u}_h^n; \mathbf{u}_h^n, \mathbf{v}_h) - b(\mathbf{v}_h, p_h^n) \quad \forall \mathbf{v}_h \in \mathbf{V}_h.$$

Step 4. $n \to n+1$. Go to Step 2 and 3 until $\|\mathbf{u}_h^{(n)} - \mathbf{u}_h^{(n-1)}\|_V < \varepsilon_{\text{tol}}$.

These algorithms have been applied to Stokes VI problems similar to Problems VE_h^0 and VE_h^1, with convergence analysis given in [28, 29, 49] for Algorithm 1 and in [28, 49] for Algorithm 2. Convergence of Algorithms 1 and 2 can be proved following the arguments in [28, Theorem 4.1], [29, Theorem 4.4] or [49, Theorem 3.3, Theorem 3.4] and is omitted in this paper.

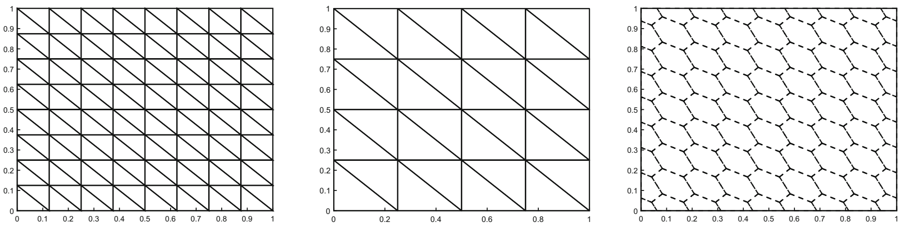
The linearization form $\widetilde{c}(\mathbf{u}_h^{n-1}; \mathbf{u}_h^{n-1}, \mathbf{v}_h)$ accompanied with convergence (cf. [23]) is utilized in the above two algorithms for solving $\mathbf{u}_h^n$. Moreover, ε_{tol} is the tolerance error and is fixed as 10^{-6} in the following tests.

6.2 Numerical Results

Let $\Omega = (0, 1)^2$, whose boundary is split into the potential slip boundary $\Gamma_S = (0, 1) \times \{0\}$ and the remaining Dirichlet boundary $\Gamma_D = \partial\Omega \setminus \Gamma_S$. We use a sequence of uniform triangular meshes for the primal partition of the domain shown in Fig. 3, as well as its corresponding dual partitions for the velocity with $h = 1/8$, respectively. We take the parameters $\nu = 1.0$ and $\rho = 10.0$ without special explanation.

For a given pair of smooth functions $(\mathbf{u}_0, p_0)$ and a fixed function g on Γ_S, it can be seen that

$$\begin{cases} g(\mathbf{x}) > |\boldsymbol{\sigma}_{\boldsymbol{\tau}}(\mathbf{u}_0)(\mathbf{x})| \text{ for all nodes } \mathbf{x} \in \Gamma_S \Rightarrow (\mathbf{u}_0, p_0) \text{ is the solution} \Rightarrow \text{No-slip occurs,} \\ g(\mathbf{x}_0) = |\boldsymbol{\sigma}_{\boldsymbol{\tau}}(\mathbf{u}_0)(\mathbf{x}_0)| \text{ for some nodes } \mathbf{x}_0 \in \Gamma_S \Rightarrow (\mathbf{u}_0, p_0) \text{ is not a solution} \Rightarrow \text{Slip occurs.} \end{cases}$$



(a) Primal partitions ($h = 1/8$) (b) Primal partitions ($h = 1/4$) (c) Dual partitions ($h = 1/8$)

Fig. 3 Case I: The primal meshes **a** for the velocity and pressure, Case II: The primal meshes **a** for the velocity and **b** for the pressure; the dual meshes **c** for the velocities

We use the numerical solution $(\mathbf{u}_{\text{ref}}, p_{\text{ref}})$ on a sufficient finer mesh ($h = 2^{-8}$) as the exact solution. We then compare the solutions $(\mathbf{u}_h, p_h)$ on the coarse meshes ($h = 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}$) with the reference solution. Since the dual meshes are not nested when the mesh is refined, which gives rise to an additional error in computing the numerical solution errors. We will compute the errors on the primal meshes according to the nature of the FVMs that we obtain solutions on the nodes of the primal subdivision.

The convergence behaviors of the finite volume approximation are evaluated as follows:

$$E_{L^2}^{\mathbf{u}}(h) = \|\mathbf{u}_h - \mathbf{u}_{\text{ref}}\|_0, \quad E_{L^2}^{p}(h) = \|p_h - p_{\text{ref}}\|_0, \quad E_{H^1}^{\mathbf{u}}(h) = |\mathbf{u}_h - \mathbf{u}_{\text{ref}}|_{H^1},$$
$$\text{order} = \ln(E_i^j(h)/E_i^j(h/2))/\ln(2), \quad i = L^2 \text{ or } H^1, \quad j = \mathbf{u} \text{ or } p.$$

We use $\mathbf{f} = -\nabla \cdot \mathbb{T}(\mathbf{u}_0, p_0)$, where

$$\mathbf{u}_0(x, y) = \begin{pmatrix} 20x^2(x-1)^2 y(y-1)(2y-1) \\ -20x(x-1)(2x-1)y^2(y-1)^2 \end{pmatrix}, \quad p_0(x, y) = 20(2x-1)(2y-1).$$

It is easy to specify $\boldsymbol{\sigma}_{\boldsymbol{\tau}}(\mathbf{u}_0)$ as follows

$$\boldsymbol{\sigma}_{\boldsymbol{\tau}}(\mathbf{u}_0) = \begin{pmatrix} 20x^2(x-1)^2 \\ 0 \end{pmatrix} \quad \text{on} \quad \Gamma_S.$$

By a direct computation, we find that

$$\max_{\Gamma_S} |\boldsymbol{\sigma}_{\boldsymbol{\tau}}(\mathbf{u}_0)| = \max_{0 \le x \le 1} |20x^2(x-1)^2| = 1.25.$$

Now for a constant g,

$$\begin{cases} g > 1.25 \Rightarrow (\mathbf{u}_0, p_0) \text{ remains the solution} \Rightarrow \text{No-slip occurs.} \\ g \le 1.25 \Rightarrow (\mathbf{u}_0, p_0) \text{ is no longer a solution} \Rightarrow \text{Slip occurs.} \end{cases}$$

Similar examples have been used in the context of inequality problems for the Stokes and Navier–Stokes equations in [24, 29]. The numerical results in Tables 1. 2, 3 and 4 and Fig. 4 perform to agree with the solution's desired behavior as indicated in Sect. 5, and the order in the last row of each table reads as the average value of the three convergence rates in each column. In Fig. 4, we draw the tangential components $\sigma_{\boldsymbol{\tau}}$ and u_τ on the x-axis along the slip boundary under Algorithm 2 for Case I, we can distinctly observe the relations between the occurrence of slip and friction function g. The results of Case II are similar to those presented in Fig. 4 because both two cases adopt the same materials of the mathematical model, and they won't be repeated. The fast convergence behaviors are revealed in Table 5 of the two iterative algorithms.

Table 1 Convergence behaviors for the Case I with projection stabilizer (Algorithm 1)

h	$g=0.1$ $(\rho=100)$			$g=0.8$ $(\rho=10)$			$g=2$ $(\rho=10)$		
	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$
2^{-3}	1.48e−01	1.02	1.23e−01	1.38e−01	1.06	6.48e−03	1.37e−01	1.12	2.37
2^{-4}	5.23e−02	4.09e−01	3.91e−02	5.09e−02	4.44e−01	3.84e−02	5.17e−02	4.74e−01	9.35e−01
2^{-5}	1.50e−02	1.42e−01	3.16e−03	1.49e−02	1.57e−01	3.18e−03	1.54e−02	1.71e−01	3.15e−01
2^{-6}	3.82e−03	4.54e−02	1.42e−03	3.81e−03	4.95e−02	1.43e−03	4.13e−03	6.01e−02	9.89e−02
Order	1.98	1.65	1.15	1.97	1.66	1.16	1.89	1.51	1.67

Table 2 Convergence behaviors for the Case I with projection stabilizer (Algorithm 2)

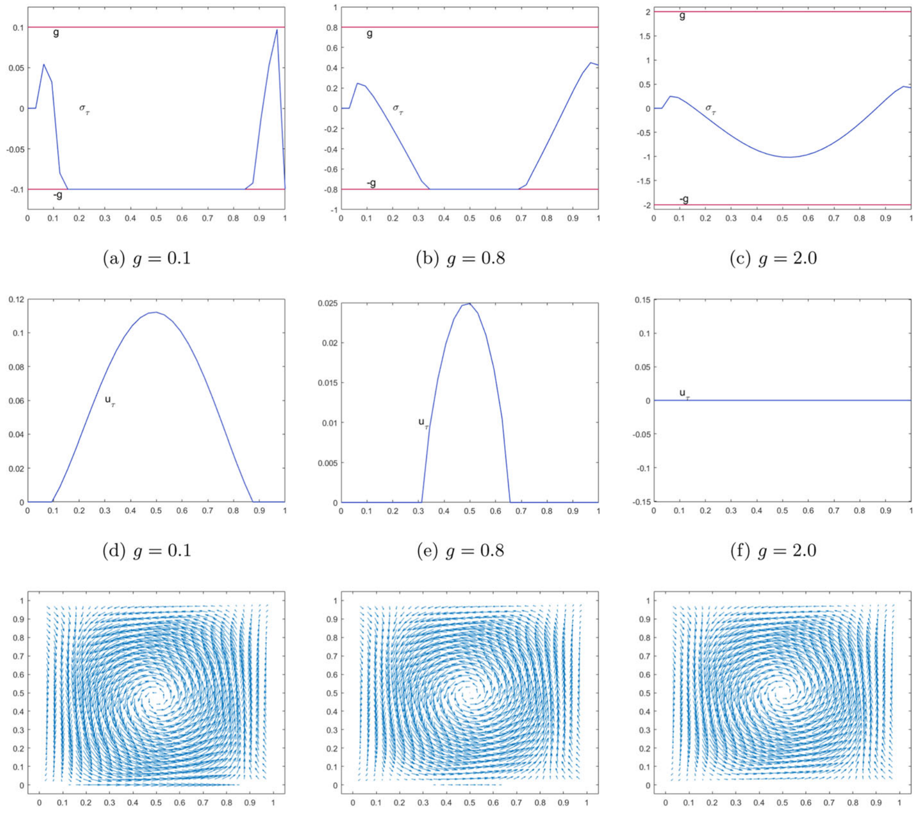
h	$g=0.1$ $(\rho=6)$			$g=0.8$ $(\rho=10)$			$g=2$ $(\rho=10)$		
	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$
2^{-3}	1.36e−01	1.02	4.28e−03	1.36e−01	1.09	8.51e−03	1.60e−01	1.13	2.03
2^{-4}	5.29e−02	4.11e−01	3.92e−02	5.09e−02	4.50e−01	3.83e−02	5.17e−02	4.74e−01	9.35e−01
2^{-5}	1.51e−02	1.43e−01	3.16e−03	1.50e−02	1.59e−01	3.15e−03	1.54e−02	1.71e−01	3.15e−01
2^{-6}	3.83e−03	4.54e−02	1.41e−03	3.83e−03	5.10e−02	1.43e−03	4.13e−03	6.01e−02	9.89e−02
Order	1.98	1.65	1.16	1.97	1.64	1.14	1.89	1.51	1.67

Table 3 Convergence behaviors for the Case II (Algorithm 1)

h	$g=0.1$ $(\rho=20)$			$g=0.8$ $(\rho=20)$			$g=2$ $(\rho=6)$		
	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$
2^{-3}	7.58e−02	1.19	2.69e−01	7.10e−02	1.16	2.66e−01	6.84e−02	1.15	5.83e−01
2^{-4}	2.40e−02	7.01e−01	3.55e−02	2.34e−02	7.06e−01	3.59e−02	2.29e−02	7.04e−01	2.42e−01
2^{-5}	6.68e−03	3.77e−01	1.20e−02	6.61e−03	3.85e−01	1.20e−02	6.58e−03	3.86e−01	8.44e−02
2^{-6}	1.75e−03	1.95e−02	2.16e−03	1.75e−03	2.00e−01	2.16e−03	1.76e−03	2.02e−01	2.80e−02
Order	1.93	0.95	2.48	1.92	0.94	2.48	1.90	0.94	1.59

Table 4 Convergence behaviors for the Case II (Algorithm 2)

h	$g=0.1$ $(\rho=6)$			$g=0.8$ $(\rho=6)$			$g=2$ $(\rho=6)$		
	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$	$E^{\boldsymbol{u}}_{L^2}$	$E^{\boldsymbol{u}}_{H^1}$	$E^{p}_{L^2}$
2^{-3}	6.81e−02	1.12	2.66e−01	6.91e−02	1.14	2.64e−01	6.65e−02	1.14	5.70e−01
2^{-4}	2.34e−02	6.94e−01	3.55e−02	2.49e−02	7.34e−01	3.55e−02	2.28e−02	7.03e−01	2.37e−01
2^{-5}	6.65e−03	3.77e−01	1.20e−02	6.90e−03	3.94e−01	1.20e−02	6.58e−03	3.86e−01	8.44e−02
2^{-6}	1.75e−03	1.95e−02	2.16e−03	1.81e−03	2.04e−01	2.16e−03	1.76e−03	2.02e−01	2.80e−02
Order	1.92	0.95	2.48	1.93	0.95	2.48	1.90	0.94	1.59



(a) $g = 0.1$ (b) $g = 0.8$ (c) $g = 2.0$

(d) $g = 0.1$ (e) $g = 0.8$ (f) $g = 2.0$

(g) $g = 0.1$ (h) $g = 0.8$ (i) $g = 2.0$

Fig. 4 Plots of $\sigma_{\boldsymbol{\tau}}$, u_τ along Γ_S and velocity fields in Ω (Case I) under Algorithm 2

Table 5 Numbers of iteration required for Algorithms 1 and 2 to converge

$h = 1/32$	Case I			Case II		
	$g = 0.1$	$g = 0.8$	$g = 2.0$	$g = 0.1$	$g = 0.8$	$g = 2.0$
Algorithm 1	33	75	63	25	50	69
Algorithm 2	5	7	6	6	7	6

7 Conclusions

In this work, we study two lowest-order conforming FVMs to solve a Navier-Stokes variational inequality. We have shown the equivalence of the numerical schemes between FVM and FEM, the discrete inf-sup conditions, existence and uniqueness of finite volume solutions. We have also derived an optimal order error estimate. It will be interesting to study the numerical solutions of the stationary Navier-Stokes equations with more complicated friction boundary conditions in the future.

Author Contributions All authors contributed to the study conception and algorithms design. The manuscript was written by Feifei Jing and Wenjing Yan and all authors commented on the present version of the manuscript.

Data Availability This manuscript has no associated data.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Adams, R.A.: Sobolev Spaces. Academic Press, New York (1975)
2. Amara, M., Dabaghi, F.: An optimal C0 finite element algorithm for the 2D biharmonic problem: theoretical analysis and numerical results. Numer. Math. **90**, 19–46 (2001)
3. Atkinson, K., Han, W.: Theoretical Numerical Analysis: A Functional Analysis Framework, 3rd edn. Springer, New York (2009)
4. Bochev, P.B., Dohrmann, C.R., Gunzburger, M.D.: Stabilization of low-order mixed finite element for the Stokes equations. SIAM J. Numer. Anal. **44**, 82–101 (2006)
5. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Springer, New York (2008)
6. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer Series in Computational Mathematics, vol. 15. Springer, New York (1991)
7. Cai, Z.Q.: On the finite volume element method. Numer. Math. **58**, 713–735 (1991)
8. Carstensen, C., Dond, A.K., Nataraj, N., Pani, A.K.: Three first-order finite volume element methods for Stokes equations under minimal regularity assumptions. SIAM J. Numer. Anal. **56**(4), 2648–2671 (2018)
9. Chou, S.H.: Analysis and convergence of a covolume method for the generalized Stokes problem. Math. Comput. **66**(217), 85–104 (1997)
10. Chou, S.H., Kwak, D.Y., Vassilevski, P.S.: Mixed covolume methods for elliptic problems on triangular grids. SIAM J. Numer. Anal. **35**(5), 1850–1861 (1998)
11. Chou, S.H., Ye, X.: Unified analysis of finite volume methods for second order elliptic problems. SIAM J. Numer. Anal. **45**(4), 1639–1653 (2007)
12. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
13. Cui, M., Ye, X.: Unified analysis of finite volume methods for the Stokes equations. SIAM J. Numer. Anal. **48**(3), 824–839 (2010)
14. Djoko, J.K., Koko, J.: Numerical methods for the Stokes and Navier–Stokes equations driven by threhold slip boundary conditions. Comput. Methods Appl. Mech. Eng. **305**, 936–958 (2016)
15. Franca, L.P., Stenberg, R.: Error analysis of some Galerkin least squares methods for the elasticity equations. SIAM J. Numer. Anal. **28**, 1680–1697 (1991)
16. Fujita, H.: Flow Problems with Unilateral Boundary Conditions. College de France, Lecons (1993)
17. Fujita, H.: A mathematical analysis of motions of viscous incompressible fluid under leak or slip boundary conditions. RIMS Kokyuroku **88**, 199–216 (1994)
18. Fujita, H.: A coherent analysis of Stokes flows under boundary conditions of friction type. J. Comput. Appl. Math. **149**, 57–69 (2002)
19. Fujita, H., Kawarada, H., Sasamoto, A.: Analysis and numerical approaches to stationary flow problms wth leak and slip boundary conditions. Lecture Notes in Numer. Appl. Anal. 14, Kinokuniya, Tokyo, pp. 17–31 (1995)
20. Girault, V., Raviart, P.A.: Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms, volume 5 of Springer Series in Computational Mathematics. Springer, Berlin (1986)
21. Glowinski, R.: Numerical Methods of Nonlinear Variational Problems. Springer, New-York (1984)
22. Glowinski, R., Lions, J., Tremolieres, R.: Numerical Analysis of Variational Inequalities. Elsevier Science Ltd, Amsterdam (1981)
23. He, Y.N., Li, J.: Convergence of three iterative methods based on the finite element discretization for the stationary Navier–Stokes equations. Comput. Math. Appl. Mech. Eng. **198**(15–16), 1351–1359 (2009)
24. Jing, F.F., Han, W., Yan, W.J., Wang, F.: Discontinuous Galerkin finite element methods for stationary Navier–Stokes problem with a nonlinear slip boundary condition of friction type. J. Sci. Comput. **76**(2), 888–912 (2018)
25. Jing, F.F., Kashiwabara, T., Yan, W.J.: Numerical analysis of the lowest-order finite volume methods for a class of variational inequality problems arising in hydrodynamics. *Submitted*
26. Jing, F.F., Li, J., Chen, Z.: Stabilized finite element methods for a blood flow model of arterosclerosis. Numer. Methods Partial Differ. Equ. **31**, 2063–2079 (2015)

27. Jing, F.F., Li, J., Chen, Z., Zhang, Z.H.: Numerical analysis of a characteristic stabilized finite element method for the time-dependent Navier–Stokes equations with nonlinear slip boundary conditions. J. Comput. Appl. Math. **320**, 43–60 (2017)
28. Kashiwabara, T.: Finite element method for Stokes equations under leak boundary condition of friction type. SIAM J. Numer. Anal. **51**(4), 2445–2469 (2013)
29. Kashiwabara, T.: On a finite element approximation of the Stokes equations under a slip boundary condition of the friction type. Jpn. J. Ind. Appl. Math. **30**, 227–261 (2013)
30. Kashiwabara, T.: On a strong solution of the non-stationary Navier–Stokes equations under slip or leak boundary conditions of friction type. J. Differ. Equ. **254**, 756–778 (2013)
31. Li, J., Chen, Z.: Optimal and analysis of finite volume methods for the stationary Navier–Stokes equations with large data. Numer. Math. **126**(1), 75–101 (2014)
32. Li, J., He, Y.N.: A stabilized finite element method based on two local Gauss integrations for the Stokes equations. J. Comput. Appl. Math. **214**(1), 58–65 (2008)
33. Li, J., Shen, L.H., Chen, Z.: Convergence and stability of a stabilized finite volume method for the stationary Navier–Stokes equations. BIT **50**(4), 823–842 (2010)
34. Li, R.H.: Generalized difference methods for a nonlinear Dirichlet problem. SIAM J. Numer. Anal. **24**(1), 77–88 (1987)
35. Li, Y., Li, K.T.: Locally stabilized finite element method for the Stokes problem with nonlinear slip boundary conditions. J. Comput. Math. **28**(6), 826–836 (2010)
36. Li, Y., Li, K.T.: Existence of the solution to stationary Navier–Stokes equations with nonlinear slip boundary conditions. J. Math. Anal. Appl. **381**, 1–9 (2011)
37. Li, Y., Li, K.T.: Uzawa iteration method for Stokes type variational inequality of the second kind. Acta Math. Appl. Sin. **17**, 303–316 (2011)
38. Lin, Y.P., Liu, J.G., Yang, M.: Finite volume element methods: an overview on recent developments. Int. J. Numer. Anal. Model., Ser. B **4**, 14–34 (2013)
39. Ling, M., Han, W., Zeng, S.D.: A pressure projection stabilized mixed finite element method for a Stokes hemivariational inequality. J. Sci. Comput. **92**, 1–21 (2022)
40. Mu, L., Ye, X.: A finite volume method for solving Navier–Stokes problems. Nonlinear Anal. Theor. Meth. Appl. **74**(17), 6686–6695 (2011)
41. Quarteroni, A., Ruiz-Baier, R.: Analysis of a finite volume element method for the Stokes problem. Numer. Math. **118**, 737–64 (2011)
42. Saito, N.: On the Stokes equations with the leak and slip boundary conditions of friction type: regularity of solutions. Pub RIMS **40**, 345–383 (2004)
43. Sheng, Y., Zhang, T., Jiang, Z.Z.: A stabilized finite volume method for stationary Naiver–Stokes equations. Chaos, Solitons Fractals **89**, 363–372 (2016)
44. Stadler, G.: Semismooth Newton and augmented Lagrangean methods for a simplified friction problem. SIAM J. Optim. **15**, 39–62 (2004)
45. Temam, R.: Navier–Stokes Equations. Theory and Numerical Analysis., 3rd edn. North-Holland, Amsterdam (1983)
46. Wang, J.P., Wang, Y.Q., Ye, X.: A new finite volume method for the Stokes problems. Int. J. Numer. Anal. Model. **7**(2), 281–302 (2010)
47. Ye, X.: On the relationship between finite volume and finite element methods applied to the Stokes equations. Numer. Methods Partial Differ. Equ. **17**(5), 440–453 (2001)
48. Zhang, T., Tang, L.X.: A stabilized finite volume method for Stokes equations using the lowest order P1–P0 element pair. Adv. Comput. Math. **41**, 781–798 (2015)
49. Zhou, G.Y., Jing, F.F., Kashiwabara, T.: The numerical methods for the coupled fluid flow under the leak interface condition of the friction-type. Numer. Math. **153**(4), 729–773 (2023)

Authors and Affiliations

Feifei Jing[1] · Weimin Han[2] · Takahito Kashiwabara[3] · Wenjing Yan[4]

Feifei Jing
ffjing@nwpu.edu.cn

Weimin Han
weimin-han@uiowa.edu

Takahito Kashiwabara
tkashiwa@ms.u-tokyo.ac.jp

1 School of Mathematics and Statistics and Xi'an Key Laboratory of Scientific Computation and Applied Statistics, Northwestern Polytechnical University, Xi'an 710129, Shaanxi, China

2 Department of Mathematics, University of Iowa, Iowa City, IA 52242-1410, USA

3 The Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro 153-8914, Japan

4 School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China