J Sci Comput
DOI 10.1007/s10915-015-0008-5

Analysis of an Augmented HDG Method for a Class of Quasi-Newtonian Stokes Flows

Gabriel N. Gatica[1] · **Filánder A. Sequeira**[2,3]

Received: 27 August 2014 / Revised: 28 January 2015 / Accepted: 5 March 2015

Abstract In this paper we introduce and analyze a hybridizable discontinuous Galerkin (HDG) method for numerically solving a class of nonlinear Stokes models arising in quasi-Newtonian fluids. Similarly as in previous papers dealing with the application of mixed finite element methods to these nonlinear models, we use the incompressibility condition to eliminate the pressure, and set the velocity gradient as an auxiliary unknown. In addition, we enrich the HDG formulation with two suitable augmented equations, which allows us to apply known results from nonlinear functional analysis, namely a nonlinear version of Babuška–Brezzi theory and the classical Banach fixed-point theorem, to prove that the discrete scheme is well-posed and derive the corresponding a priori error estimates. Then we discuss some general aspects concerning the computational implementation of the method, which show a significant reduction of the size of the linear systems involved in the Newton iterations. Finally, we provide several numerical results illustrating the good performance of the proposed scheme and confirming the optimal order of convergence provided by the HDG approximation.

Keywords Nonlinear Stokes model · Mixed finite element method · Hybridized discontinuous Galerkin method · Augmented formulation

This work was partially supported by CONICYT-Chile through BASAL project CMM, Universidad de Chile, project Anillo ACT1118 (ANANUM), and the Becas-Chile Programme for foreign students; and by Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción.

✉ Gabriel N. Gatica
ggatica@ci2ma.udec.cl

Filánder A. Sequeira
filander.sequeira@una.cr; fsequeira@ci2ma.udec.cl

[1] Centro de Investigación en Ingeniería Matemática, Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile

[2] Escuela de Matemática, Universidad Nacional de Costa Rica, Heredia, Costa Rica

[3] Present Address: Centro de Investigación en Ingeniería Matemática, Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile

Published online: 15 March 2015

1 Introduction

The devising of suitable numerical methods for solving the linear and nonlinear Stokes and related problems has become a very active research area during the last decade. In particular, a mixed finite element method and a suitable augmented version of the latter for a nonlinear Stokes flow problem involving a non-Newtonian fluid, are introduced and analized in [21]. In addition, the velocity–pressure–stress formulation for incompressible flows has gained considerable attention in recent years due to its natural applicability to non-Newtonian flows, where the corresponding constitutive equations are nonlinear. In general, an interesting feature of the mixed methods is given by the fact that, besides the original unknowns, they yield direct approximations of several other quantities of physical interest. For instance, an accurate direct calculation of the stresses is very desirable for flow problems involving interaction with solid structures.

On the other hand, the hybridizable discontinuous Galerkin (HDG) method, introduced in [10] for diffusion problems, is one of the several high-order discretization schemes that benefit from the hybridization technique originally applied in [15] to the local discontinuous Galerkin (LDG) method for time dependent convection–diffusion problems. The main advantages of HDG methods include a substantial reduction of the globally coupled degrees of freedom, which was a criticism for the discontinuous Galerkin (DG) methods for elliptic problems during the last decade, and the fact that convergence is obtained even for a polynomial degree $k = 0$. Additionally, the approximate flux converges with order $k + 1$ for $k \geq 0$, and an element-by-element computation of a new approximation of the scalar variable is possible, which converges with order $k + 2$ for $k \geq 1$ (see e.g. [9,11,13]). In the context of the linear Stokes equation, the hybridization for DG methods was initially introduced in [5] and then analyzed in [11,30]. Lately, an overview of the recent work by Cockburn et al. on the devising of hybridizable discontinuous Galerkin (HDG) methods for the Stokes equations of incompressible flow was provided in [14].

Now, the utilization of DG methods to numerically solve nonlinear boundary value problems has been first considered in [3] and [24]. Indeed, the application of the local discontinuous Galerkin (LDG) method to a class of nonlinear diffusion problems was developed in [3], whereas the extension of the interior penalty *hp* DG method to quasilinear elliptic equations was studied in [24]. The results from [3] were generalized in [4], where the a-priori and a-posteriori error analyses of the LDG method as applied to certain type of nonlinear Stokes models (whose kinematic viscosities are nonlinear monotone functions of the gradient of the velocity) were derived. The approach in [4] is based on the introduction of the flux and the tensor gradient of the velocity as further unknowns. A suitable Lagrange multiplier is also needed to ensure that the corresponding discrete variational formulation is well-posed. A two-fold saddle point operator equation is obtained as the resulting LDG mixed formulation, which is then reduced to a dual mixed formulation. A nonlinear version of the well known Babuška–Brezzi theory is applied to prove that the discrete formulation is well-posed and derive the corresponding a priori error analysis. In turn, the analysis from [24] was extended in [16], where the a priori and a posteriori error analysis, with respect to a mesh-dependent energy norm, of a class of interior penalty *hp* DGFEM for the numerical approximation of basically the same quasi-Newtonian fluid flow problems studied in [4], were provided. Furthermore, an HDG approach was employed in [29] for the numerical solution of steady and time-dependent nonlinear convection–diffusion equations. In fact, the approximate scalar variable and corresponding flux are first expressed in [29] in terms of an approximate trace of the scalar variable, and then the jump condition of the numerical fluxes are explicitly enforced

across the element boundaries. As a consequence, a global equation system solely in terms of the approximate trace of the scalar variable is obtained at every Newton iteration. At the end, and similarly as in previous papers on HDG, an element-by-element postprocessing scheme is applied to obtain new approximations of the flux and the scalar variable, which converge with order $k+1$ and $k+2$, respectively, in the L^2-norm. Nevertheless, and up to our knowledge, there is still no contribution in the literature concerning HDG for nonlinear Stokes systems.

According to the above discussion, we are interested in this paper in applying the HDG approach to the class of quasi-Newtonian Stokes flows studied in [4,16,19] (see also [21,25]). To this end, we plan to employ the same velocity–pseudostress formulation from [21]. In what follows, given any Hilbert space U, $\mathbf{U} := U^n$ and $\mathbb{U} := U^{n\times n}$ denote, respectively, the space of vector and square matrices of order n, $n \in \{2, 3\}$, with entries in U. In order to define the boundary value problem of interest, we now let Ω be a bounded and simply connected polygonal domain in R^n with boundary Γ. As in [21], our goal is to determine the velocity $\mathbf{u}$, the pseudostress tensor $\boldsymbol{\sigma}$, and the pressure p of a steady flow occupying the region Ω, under the action of external forces. More precisely, given a volume force $\mathbf{f} \in \mathbf{L}^2(\Omega)$ and $\mathbf{g} \in \mathbf{H}^{1/2}(\Gamma)$, we seek a tensor field $\boldsymbol{\sigma}$, a vector field $\mathbf{u}$, and a scalar field p such that

$$\begin{aligned} \boldsymbol{\sigma} &= \mu(|\nabla\mathbf{u}|)\,\nabla\mathbf{u} - p\,\mathbb{I} \quad \text{in} \quad \Omega, \qquad \mathbf{div}(\boldsymbol{\sigma}) = -\mathbf{f} \quad \text{in} \quad \Omega, \\ \mathrm{div}(\mathbf{u}) &= 0 \quad \text{in} \quad \Omega, \qquad \mathbf{u} = \mathbf{g} \quad \text{on} \quad \Gamma, \qquad \int_\Omega p = 0, \end{aligned} \tag{1.1}$$

where $\mu : R^+ \to R^+$ is the nonlinear kinematic velocity function of the fluids, **div** stands for the usual divergence operator div acting along each row of tensor, $\nabla\mathbf{u}$ is the tensor gradient of $\mathbf{u}$, $|\cdot|$ is the euclidean norm of $R^{n\times n}$, and $\mathbb{I}$ is the identity matrix of $R^{n\times n}$. As required by the incompressibility condition, we assume from now on that the datum $\mathbf{g}$ satisfies the compatibility condition $\int_\Gamma \mathbf{g}\cdot\boldsymbol{\nu} = 0$, where $\boldsymbol{\nu}$ stands for the unit outward normal at Γ. The kind of nonlinear Stokes problem given by (1.1) appears in the modeling of a large class of non-Newtonian fluids (see, e.g. [1,27,28,32]). In particular, the Ladyzhenskaya law, is given by $\mu(t) := \mu_0 + \mu_1 t^{\beta-2}\ \forall\, t \in R^+$, with $\mu_0 \geq 0$, $\mu_1 > 0$, and $\beta > 1$, and the Carreau law for viscoplastic flows (see, e.g. [28,32]) reads $\mu(t) := \mu_0 + \mu_1(1+t^2)^{(\beta-2)/2}\ \forall\, t \in R^+$, with $\mu_0 \geq 0$, $\mu_1 > 0$, and $\beta \geq 1$.

The rest of the work is organized as follows. In Sect. 2 we introduce the augmented hybridizable discontinuous Galerkin formulation involving the velocity, the pseudostress, the velocity gradient and the trace of the velocity, as unknowns. In Sect. 3 we show the unique solvability of the augmented HDG scheme by considering an equivalent formulation and then applying a nonlinear version of the Babuška–Brezzi theory and the classical Banach fixed-point Theorem. The corresponding a priori error estimates are derived in Sect. 4. Next, in Sect. 5 we discuss some general aspects concerning the computational implementation of the HDG method. Finally, several numerical experiments validating the good performance of the method and confirming the rates of convergence derived are reported in Sect. 6. We end the present section with further notations to be used below. Given $\boldsymbol{\tau} := (\tau_{ij})$, $\boldsymbol{\zeta} := (\zeta_{ij}) \in R^{n\times n}$, we write as usual

$$\mathrm{tr}\,(\boldsymbol{\tau}) := \sum_{i=1}^{n} \tau_{ii}, \quad \boldsymbol{\tau}^{\mathrm{d}} := \boldsymbol{\tau} - \frac{1}{n}\mathrm{tr}\,(\boldsymbol{\tau})\,\mathbb{I}, \quad \text{and} \quad \boldsymbol{\tau} : \boldsymbol{\zeta} := \sum_{i,j=1}^{n} \tau_{ij}\zeta_{ij}.$$

Also, in what follows we utilize the standard terminology for Sobolev spaces and norms, employ $\mathbf{0}$ to denote a generic null vector, null tensor or null operator, and use C, with

or without subscripts, bars, tildes or hats, to denote generic constants independent of the discretization parameters, which may take different values at different places.

2 The Augmented HDG Method

2.1 The Hybridizable Discontinuous Galerkin Method

We begin by eliminating the pressure. Indeed, we know from [21, Section 2.1] that the pair given by the first and third equations in (1.1) is equivalent to

$$\boldsymbol{\sigma} \;=\; \mu(|\nabla \mathbf{u}|)\nabla \mathbf{u} - p\,\mathbb{I} \;\text{ in } \Omega \;\text{ and }\; p \;=\; -\frac{1}{n}\,\mathrm{tr}\,(\boldsymbol{\sigma}) \;\text{ in } \Omega. \tag{2.1}$$

In what follows we let $\psi_{ij} : R^{n\times n} \to R$ be the mapping given by $\psi_{ij}(\mathbf{r}) := \mu(|\mathbf{r}|)r_{ij}$ for all $\mathbf{r} := (r_{ij}) \in R^{n\times n}$, for all $i, j \in \{1, \ldots, n\}$. Then, throughout this paper we assume that μ is of class C^1 and that there exist $\gamma_0, \alpha_0 > 0$ such that for all $\mathbf{r} := (r_{ij}), \mathbf{s} := (s_{ij}) \in R^{n\times n}$, there holds

$$|\psi_{ij}(\mathbf{r})| \;\le\; \gamma_0 \|\mathbf{r}\|_{R^{n\times n}}, \quad \left|\frac{\partial}{\partial r_{kl}}\psi_{ij}(\mathbf{r})\right| \;\le\; \gamma_0, \quad \forall \quad i, j, k, l \in \{1, \ldots, n\}, \tag{2.2}$$

and

$$\sum_{i,j,k,l=1}^{n} \frac{\partial}{\partial r_{kl}}\psi_{ij}(\mathbf{r}) s_{ij} s_{kl} \;\ge\; \alpha_0 \|\mathbf{s}\|^2_{R^{n\times n}}. \tag{2.3}$$

It is easy to check that the Carreau law satisfies (2.2) and (2.3) for all $\mu_0 > 0$, and for all $\beta \in [1, 2]$. In particular, with $\beta = 2$ we recover the usual linear Stokes model. We observe in advance that the above assumptions are required to prove later on the strong monotonicity and Lipschitz-continuity properties of the continuous and discrete nonlinear operators involving the viscosity function μ (see Lemmas 3.4, 3.5 and 3.6 below).

Observe that we can rewrite (2.1) as

$$\boldsymbol{\sigma} \;=\; \boldsymbol{\psi}(\nabla \mathbf{u}) - p\,\mathbb{I} \;\text{ in } \Omega \;\text{ and }\; p \;=\; -\frac{1}{n}\mathrm{tr}\,(\boldsymbol{\sigma}) \;\text{ in } \Omega,$$

where $\boldsymbol{\psi} : R^{n\times n} \to R^{n\times n}$ is given by $\boldsymbol{\psi}(\mathbf{r}) := (\psi_{ij}(\mathbf{r}))$ for all $\mathbf{r} := (r_{ij}) \in R^{n\times n}$. Hence, replacing p by $-\frac{1}{n}\mathrm{tr}\,(\boldsymbol{\sigma})$ in the first equation of (1.1), and introducing the gradient $\mathbf{t} := \nabla \mathbf{u}$ in Ω as an auxiliary unknown, we arrive at the system

$$\begin{array}{rlrl} \boldsymbol{\psi}(\mathbf{t}) - \boldsymbol{\sigma}^{\mathrm{d}} &= \mathbf{0} \;\text{ in }\; \Omega, & \mathbf{t} - \nabla \mathbf{u} &= \mathbf{0} \;\text{ in }\; \Omega, \\ -\mathbf{div}(\boldsymbol{\sigma}) &= \mathbf{f} \;\text{ in }\; \Omega, & \mathrm{tr}\,(\mathbf{t}) &= 0 \;\text{ in }\; \Omega, \\ \mathbf{u} &= \mathbf{g} \;\text{ on }\; \Gamma, & \displaystyle\int_\Omega \mathrm{tr}\,(\boldsymbol{\sigma}) &= 0. \end{array} \tag{2.4}$$

We recall here that a well-posed continuous formulation of (2.4) has been proposed in [21, Section 2], which reads: Find $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in X_1 \times M_1 \times \mathbf{L}^2(\Omega)$ such that

$$\begin{array}{rl} \displaystyle\int_\Omega \boldsymbol{\psi}(\mathbf{t}) : \mathbf{s} - \int_\Omega \boldsymbol{\sigma}^{\mathrm{d}} : \mathbf{s} = 0 & \forall\, \mathbf{s} \in X_1, \\ \displaystyle -\int_\Omega \mathbf{t} : \boldsymbol{\tau}^{\mathrm{d}} - \int_\Omega \mathbf{u} \cdot \mathbf{div}(\boldsymbol{\tau}) = -\langle \boldsymbol{\tau}\boldsymbol{\nu}, \mathbf{g}\rangle_\Gamma & \forall\, \boldsymbol{\tau} \in M_1, \\ \displaystyle -\int_\Omega \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\sigma}) = \int_\Omega \mathbf{f} \cdot \mathbf{v} & \forall\, \mathbf{v} \in \mathbf{L}^2(\Omega), \end{array} \tag{2.5}$$

where $X_1 := \{\mathbf{s} \in \mathbb{L}^2(\Omega) \ : \ \mathrm{tr}\,(\mathbf{s}) = 0\}$ and $M_1 = \left\{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega) \ : \ \int_\Omega \mathrm{tr}\,(\boldsymbol{\tau}) = 0\right\}$. The purpose of reminding here (2.5) will become clear in the a priori error analysis given below in Section 4.

Next, in order to introduce the HDG method for the system (2.4), we first need some preliminary notations. Let $\mathcal{T}_h$ be a shape-regular triangulation of $\bar{\Omega}$ without the presence of hanging nodes, and let $\mathcal{E}_h$ be the set of faces F of $\mathcal{T}_h$. Then, we set

$$\partial\mathcal{T}_h \ := \ \cup\{\partial T \ : \ T \in \mathcal{T}_h\}\,,$$

and introduce the inner products:

$$(\mathbf{u},\mathbf{v})_{\mathcal{T}_h} := \sum_{T\in\mathcal{T}_h}\int_T \mathbf{u}\cdot\mathbf{v} \quad \forall\,\mathbf{u},\mathbf{v}\in\mathbf{L}^2(\mathcal{T}_h),$$

$$(\boldsymbol{\sigma},\boldsymbol{\tau})_{\mathcal{T}_h} := \sum_{T\in\mathcal{T}_h}\int_T \boldsymbol{\sigma}:\boldsymbol{\tau} \quad \forall\,\boldsymbol{\sigma},\boldsymbol{\tau}\in\mathbb{L}^2(\mathcal{T}_h),$$

$$\langle\mathbf{u},\mathbf{v}\rangle_{\partial\mathcal{T}_h} := \sum_{T\in\mathcal{T}_h}\int_{\partial T} \mathbf{u}\cdot\mathbf{v} \quad \forall\,\mathbf{u},\mathbf{v}\in\mathbf{L}^2(\partial\mathcal{T}_h),$$

$$\langle\mathbf{u},\mathbf{v}\rangle_{\partial\mathcal{T}_h\setminus\Gamma} := \sum_{T\in\mathcal{T}_h}\sum_{F\in\partial T\setminus\Gamma}\int_F \mathbf{u}\cdot\mathbf{v} \quad \forall\,\mathbf{u},\mathbf{v}\in\mathbf{L}^2(\partial\mathcal{T}_h),$$

with the induced norm

$$\|\mathbf{v}\|_{\mathcal{T}_h} \ := \ (\mathbf{v},\mathbf{v})_{\mathcal{T}_h}^{1/2} \quad \forall\,\mathbf{v}\in\mathbf{L}^2(\mathcal{T}_h).$$

In addition, we let $P_k(U)$ be the space of polynomials of total degree at most k defined on the domain U, and denote by $\mathcal{E}_h^i$ and $\mathcal{E}_h^\partial$ the set of interior and boundary faces, respectively, of $\mathcal{E}_h$.

On the other hand, let $\boldsymbol{\nu}^+$ and $\boldsymbol{\nu}^-$ be the outward unit normal vectors on the boundaries of two neighboring elements T^+ and T^-, respectively. We use $(\boldsymbol{\tau}^\pm,\mathbf{v}^\pm)$ to denote the traces of $(\boldsymbol{\tau},\mathbf{v})$ on $F := \partial T^+\cap\partial T^-$ from the interior of $T^\pm$, where $\boldsymbol{\tau}$ and $\mathbf{v}$ are second-order tensorial and vectorial functions, respectively. Then, we define the means $\{\!\{\cdot\}\!\}$ and jumps $[\![\cdot]\!]$ for $F\in\mathcal{E}_h^i$, as follows

$$\{\!\{\boldsymbol{\tau}\}\!\} := \frac{1}{2}\left(\boldsymbol{\tau}^+ + \boldsymbol{\tau}^-\right), \qquad \{\!\{\mathbf{v}\}\!\} := \frac{1}{2}\left(\mathbf{v}^+ + \mathbf{v}^-\right),$$
$$[\![\boldsymbol{\tau}]\!] := \boldsymbol{\tau}^+\boldsymbol{\nu}^+ + \boldsymbol{\tau}^-\boldsymbol{\nu}^-, \qquad [\![\mathbf{v}]\!] := \mathbf{v}^+\otimes\boldsymbol{\nu}^+ + \mathbf{v}^-\otimes\boldsymbol{\nu}^-,$$

where $\otimes$ denotes the usual dyadic or tensor product. Next, given $k\geq 1$, the finite dimensional discontinuous subspaces are given by

$$S_h := \left\{\mathbf{s}\in\mathbb{L}^2(\Omega) \ : \ \mathbf{s}|_T\in\mathbb{P}_k(T) \quad \forall\,T\in\mathcal{T}_h\right\},$$
$$\Sigma_h := \left\{\boldsymbol{\tau}\in\mathbb{L}^2(\Omega) \ : \ \boldsymbol{\tau}|_T\in\mathbb{P}_k(T) \quad \forall\,T\in\mathcal{T}_h, \quad \text{and} \quad \int_\Omega \mathrm{tr}\,(\boldsymbol{\tau}) = 0\right\},$$
$$V_h := \left\{\mathbf{v}\in\mathbf{L}^2(\Omega) \ : \ \mathbf{v}|_T\in\mathbf{P}_{k-1}(T) \quad \forall\,T\in\mathcal{T}_h\right\},$$
$$M_h := \left\{\boldsymbol{\mu}\in\mathbf{L}^2(\mathcal{E}_h^i) \ : \ \boldsymbol{\mu}|_F\in\mathbf{P}_k(F) \quad \forall\,F\in\mathcal{E}_h^i\right\}.$$

At this point we remark in advance that the choice of the polynomial degree $k-1$ in the definition of V_h is justified by the need of satisfying later on a joint discrete inf-sup condition with the space Σ_h (see Lemma 3.7 below).

We now proceed similarly as in [11] to derive the HDG formulation of (2.4). In fact, testing the equations in (2.4) with elements in the foregoing subspaces, integrating by parts, and introducing the numerical fluxes $\widehat{\mathbf{u}}_h$ and $\widehat{\boldsymbol{\sigma}_h\boldsymbol{\nu}}$, we arrive at: Find $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\lambda}_h) \in S_h \times \Sigma_h \times V_h \times M_h$, such that

$$(\boldsymbol{\psi}(\mathbf{t}_h), \mathbf{s}_h)_{\mathcal{T}_h} - (\mathbf{s}_h, \boldsymbol{\sigma}_h^{\mathrm{d}})_{\mathcal{T}_h} = 0 \quad \forall\, \mathbf{s}_h \in S_h, \tag{2.6a}$$

$$(\mathbf{t}_h, \boldsymbol{\tau}_h^{\mathrm{d}})_{\mathcal{T}_h} + (\mathbf{u}_h, \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} - \langle \boldsymbol{\tau}_h\boldsymbol{\nu}, \widehat{\mathbf{u}}_h\rangle_{\partial\mathcal{T}_h} = 0 \quad \forall\, \boldsymbol{\tau}_h \in \Sigma_h, \tag{2.6b}$$

$$(\boldsymbol{\sigma}_h, \nabla_h\mathbf{v}_h)_{\mathcal{T}_h} - \langle\widehat{\boldsymbol{\sigma}_h\boldsymbol{\nu}}, \mathbf{v}_h\rangle_{\partial\mathcal{T}_h} = (\mathbf{f}, \mathbf{v}_h)_{\mathcal{T}_h} \quad \forall\, \mathbf{v}_h \in V_h, \tag{2.6c}$$

$$\langle\widehat{\boldsymbol{\sigma}_h\boldsymbol{\nu}}, \boldsymbol{\mu}_h\rangle_{\partial\mathcal{T}_h\setminus\Gamma} = 0 \quad \forall\, \boldsymbol{\mu}_h \in M_h, \tag{2.6d}$$

where, letting Π_Γ be the $\mathbf{L}^2(\Gamma)$ projection onto the space of piecewise polynomials of degree less than or equals to k on $\mathcal{E}_h^\partial$, we set

$$\widehat{\mathbf{u}}_h := \begin{cases} \Pi_\Gamma(\mathbf{g}) & \text{on } \mathcal{E}_h^\partial, \\ \boldsymbol{\lambda}_h & \text{on } \mathcal{E}_h^i, \end{cases} \quad \text{and} \quad \widehat{\boldsymbol{\sigma}_h\boldsymbol{\nu}} := \boldsymbol{\sigma}_h\boldsymbol{\nu} - \mathbf{S}(\mathbf{u}_h - \widehat{\mathbf{u}}_h) \ \text{ on } \partial\mathcal{T}_h, \tag{2.6e}$$

where $\mathbf{S}$ is a stabilization operator to be defined below. Note that the condition $\widehat{\mathbf{u}}_h = \Pi_\Gamma(\mathbf{g})$ on $\mathcal{E}_h^\partial$ is usually imposed in the equivalent way $\langle\widehat{\mathbf{u}}_h, \boldsymbol{\mu}_h\rangle_\Gamma = \langle\mathbf{g}, \boldsymbol{\mu}_h\rangle_\Gamma \ \forall\, \boldsymbol{\mu}_h \in \mathbf{P}_k(\mathcal{E}_h)$, which is employed to perform the solvability analysis of (2.6). In this sense, note first that problem (2.6) can be reformulated as

$$\begin{aligned}
(\boldsymbol{\psi}(\mathbf{t}_h), \mathbf{s}_h)_{\mathcal{T}_h} - (\mathbf{s}_h, \boldsymbol{\sigma}_h^{\mathrm{d}})_{\mathcal{T}_h} &= 0, \\
(\mathbf{t}_h, \boldsymbol{\tau}_h^{\mathrm{d}})_{\mathcal{T}_h} + (\mathbf{u}_h, \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} - \langle\boldsymbol{\tau}_h\boldsymbol{\nu}, \boldsymbol{\lambda}_h\rangle_{\partial\mathcal{T}_h\setminus\Gamma} &= \langle\boldsymbol{\tau}_h\boldsymbol{\nu}, \mathbf{g}\rangle_\Gamma, \\
-(\mathbf{v}_h, \mathbf{div}_h(\boldsymbol{\sigma}_h))_{\mathcal{T}_h} + \langle\mathbf{S}(\mathbf{u}_h - \boldsymbol{\lambda}_h), \mathbf{v}_h\rangle_{\partial\mathcal{T}_h\setminus\Gamma} + \langle\mathbf{S}\mathbf{u}_h, \mathbf{v}_h\rangle_\Gamma &= (\mathbf{f}, \mathbf{v}_h)_{\mathcal{T}_h} + \langle\mathbf{S}\mathbf{g}, \mathbf{v}_h\rangle_\Gamma, \\
\langle\boldsymbol{\sigma}_h\boldsymbol{\nu}, \boldsymbol{\mu}_h\rangle_{\partial\mathcal{T}_h\setminus\Gamma} - \langle\mathbf{S}(\mathbf{u}_h - \boldsymbol{\lambda}_h), \boldsymbol{\mu}_h\rangle_{\partial\mathcal{T}_h\setminus\Gamma} &= 0,
\end{aligned}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\mu}_h) \in S_h \times \Sigma_h \times V_h \times M_h$, where (2.6c) has been rewritten using that

$$\begin{aligned}
(\boldsymbol{\sigma}_h, \nabla_h\mathbf{v}_h)_{\mathcal{T}_h} &= \sum_{T\in\mathcal{T}_h}\int_T \boldsymbol{\sigma}_h : \nabla\mathbf{v}_h \;=\; \sum_{T\in\mathcal{T}_h}\left\{-\int_T \mathbf{div}(\boldsymbol{\sigma}_h)\cdot\mathbf{v}_h + \int_{\partial T}\boldsymbol{\sigma}_h\boldsymbol{\nu}\cdot\mathbf{v}_h\right\}, \\
&= -(\mathbf{v}_h, \mathbf{div}_h(\boldsymbol{\sigma}_h))_{\mathcal{T}_h} \;+\; \langle\boldsymbol{\sigma}_h\boldsymbol{\nu}, \mathbf{v}_h\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

We complete the definition of the HDG method by describing the stabilization tensor $\mathbf{S}$. In [11], general conditions for $\mathbf{S}$ were proposed, where in particular $\mathbf{S}^+$ does not necessarily match $\mathbf{S}^-$ for each $F \in \mathcal{E}_h^i$. Here, we consider the special case in which $\mathbf{S}^+ = \mathbf{S}^-$ in each $F \in \mathcal{E}_h^i$, that is, $\mathbf{S}$ has only one value on each $F \in \mathcal{E}_h$. More precisely, given $F \in \mathcal{E}_h$, the tensor $\mathbf{S}$ satisfies the following conditions:

$$\mathbf{S}|_F \text{ is constant, and } \mathbf{S}|_F \text{ is symmetric and positive definite.}$$

Observe that $\mathbf{S}^{-1}$ is well defined and symmetric and positive definite as well on each $F \in \mathcal{E}_h$. In (3.5) below, we select a particular choice for tensor $\mathbf{S}$ in order to establish the well-posedness of (2.9).

2.2 The Augmented HDG Formulation

In order to establish the unique solvability of the nonlinear problem (2.9), we now enrich the HDG formulation with two augmented equations arising from the constitutive and equilibrium equations, that is

$$\kappa_1(\boldsymbol{\sigma}_h^{\mathrm{d}} - \boldsymbol{\psi}(\mathbf{t}_h), \boldsymbol{\tau}_h^{\mathrm{d}})_{\mathcal{T}_h} = 0 \quad \forall\, \boldsymbol{\tau}_h \in \Sigma_h, \tag{2.7}$$

and

$$\kappa_2(\mathbf{div}_h(\boldsymbol{\sigma}_h), \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} = -\kappa_2(\mathbf{f}, \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} \quad \forall\, \boldsymbol{\tau}_h \in \Sigma_h, \tag{2.8}$$

where $\kappa_1, \kappa_2 > 0$ are parameters to be determined later on. In this way, our problem becomes: Find $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\lambda}_h) \in S_h \times \Sigma_h \times V_h \times M_h$ such that

$$(\boldsymbol{\psi}(\mathbf{t}_h), \mathbf{s}_h)_{\mathcal{T}_h} - (\mathbf{s}_h, \boldsymbol{\sigma}_h^{\mathrm{d}})_{\mathcal{T}_h} = 0\,, \tag{2.9a}$$

$$(\mathbf{t}_h, \boldsymbol{\tau}_h^{\mathrm{d}})_{\mathcal{T}_h} + (\mathbf{u}_h, \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} - \langle \boldsymbol{\tau}_h \boldsymbol{\nu}, \boldsymbol{\lambda}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} = \langle \boldsymbol{\tau}_h \boldsymbol{\nu}, \mathbf{g} \rangle_{\Gamma}\,, \tag{2.9b}$$

$$-(\mathbf{v}_h, \mathbf{div}_h(\boldsymbol{\sigma}_h))_{\mathcal{T}_h} + \langle \mathbf{S}(\mathbf{u}_h - \boldsymbol{\lambda}_h), \mathbf{v}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} + \langle \mathbf{S}\mathbf{u}_h, \mathbf{v}_h \rangle_{\Gamma} = (\mathbf{f}, \mathbf{v}_h)_{\mathcal{T}_h} + \langle \mathbf{S}\mathbf{g}, \mathbf{v}_h \rangle_{\Gamma}\,, \tag{2.9c}$$

$$\langle \boldsymbol{\sigma}_h \boldsymbol{\nu}, \boldsymbol{\mu}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} - \langle \mathbf{S}(\mathbf{u}_h - \boldsymbol{\lambda}_h), \boldsymbol{\mu}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} = 0\,, \tag{2.9d}$$

$$\kappa_1(\boldsymbol{\sigma}_h^{\mathrm{d}} - \boldsymbol{\psi}(\mathbf{t}_h), \boldsymbol{\tau}_h^{\mathrm{d}})_{\mathcal{T}_h} = 0\,, \tag{2.9e}$$

$$\kappa_2(\mathbf{div}_h(\boldsymbol{\sigma}_h), \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} = -\kappa_2(\mathbf{f}, \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h}\,, \tag{2.9f}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\mu}_h) \in S_h \times \Sigma_h \times V_h \times M_h$. Hence, in what follows we proceed as in [3,4] and derive an equivalent formulation to (2.9) (see (2.11) below), for which we prove its unique solvability. In addition, the a priori error estimates for (2.9) will also be based on the analysis of (2.11). We emphasize, however, that the introduction of this equivalent formulation is just for theoretical purposes and by no means for the explicit computation of the solution of (2.9), which is solved directly as we explain below in Sect. 5.

First, we consider equation (2.9d) and note that

$$\begin{aligned}
0 &= \langle \boldsymbol{\sigma}_h \boldsymbol{\nu}, \boldsymbol{\mu}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} - \langle \mathbf{S}\mathbf{u}_h, \boldsymbol{\mu}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} + \langle \mathbf{S}\boldsymbol{\lambda}_h, \boldsymbol{\mu}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} \\
&= \sum_{T \in \mathcal{T}_h} \sum_{F \in \partial T \setminus \Gamma} \int_F \boldsymbol{\sigma}_h \boldsymbol{\nu} \cdot \boldsymbol{\mu}_h - \sum_{T \in \mathcal{T}_h} \sum_{F \in \partial T \setminus \Gamma} \int_F \mathbf{S}\mathbf{u}_h \cdot \boldsymbol{\mu}_h + \sum_{T \in \mathcal{T}_h} \sum_{F \in \partial T \setminus \Gamma} \int_F \mathbf{S}\boldsymbol{\lambda}_h \cdot \boldsymbol{\mu}_h \\
&= \sum_{F \in \mathcal{E}_h^i} \int_F [\![\boldsymbol{\sigma}_h]\!] \cdot \boldsymbol{\mu}_h - 2 \sum_{F \in \mathcal{E}_h^i} \int_F \Big(\mathbf{S}\{\!\{\mathbf{u}_h\}\!\} \cdot \boldsymbol{\mu}_h - \mathbf{S}\boldsymbol{\lambda}_h \cdot \boldsymbol{\mu}_h \Big) \\
&= \int_{\mathcal{E}_h^i} \Big([\![\boldsymbol{\sigma}_h]\!] - 2\mathbf{S}\{\!\{\mathbf{u}_h\}\!\} + 2\mathbf{S}\boldsymbol{\lambda}_h \Big) \cdot \boldsymbol{\mu}_h \quad \forall\, \boldsymbol{\mu}_h \in M_h.
\end{aligned}$$

Hence, using that $[\![\boldsymbol{\sigma}_h]\!] - 2\mathbf{S}\{\!\{\mathbf{u}_h\}\!\} + 2\mathbf{S}\boldsymbol{\lambda}_h \in M_h$, we find that

$$[\![\boldsymbol{\sigma}_h]\!] - 2\,\mathbf{S}\{\!\{\mathbf{u}_h\}\!\} + 2\,\mathbf{S}\boldsymbol{\lambda}_h = \mathbf{0} \quad \text{on } \mathcal{E}_h^i,$$

which yields

$$\boldsymbol{\lambda}_h = \{\!\{\mathbf{u}_h\}\!\} - \frac{1}{2}\mathbf{S}^{-1}[\![\boldsymbol{\sigma}_h]\!] \quad \text{on } \mathcal{E}_h^i. \tag{2.10}$$

Observe that (2.10) coincides with the expression for $\widehat{\mathbf{u}}_h$ given in [11]. We now replace $\boldsymbol{\lambda}_h$ from (2.10) in (2.9b) and (2.9c). For this purpose, we first observe that

$$\begin{aligned}
-\langle \boldsymbol{\tau}_h \boldsymbol{\nu}, \boldsymbol{\lambda}_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma} &= -\sum_{T \in \mathcal{T}_h} \sum_{F \in \partial T \setminus \Gamma} \boldsymbol{\tau}_h \boldsymbol{\nu} \cdot \boldsymbol{\lambda}_h \;=\; -\int_{\mathcal{E}_h^i} [\![\boldsymbol{\tau}_h]\!] \cdot \boldsymbol{\lambda}_h, \\
&= \frac{1}{2} \int_{\mathcal{E}_h^i} \mathbf{S}^{-1}[\![\boldsymbol{\sigma}_h]\!] \cdot [\![\boldsymbol{\tau}_h]\!] \;-\; \int_{\mathcal{E}_h^i} \{\!\{\mathbf{u}_h\}\!\} \cdot [\![\boldsymbol{\tau}_h]\!],
\end{aligned}$$

and

$$-\langle \mathbf{S}\boldsymbol{\lambda}_h, \mathbf{v}_h\rangle_{\partial\mathcal{T}_h\setminus\Gamma} = -\langle \mathbf{S}\mathbf{v}_h, \boldsymbol{\lambda}_h\rangle_{\partial\mathcal{T}_h\setminus\Gamma} = -\sum_{T\in\mathcal{T}_h}\sum_{F\in\partial T\setminus\Gamma}\mathbf{S}\mathbf{v}_h\cdot\boldsymbol{\lambda}_h,$$
$$= -2\int_{\mathcal{E}_h^i}\mathbf{S}\{\!\{\mathbf{v}_h\}\!\}\cdot\boldsymbol{\lambda}_h = \int_{\mathcal{E}_h^i}\{\!\{\mathbf{v}_h\}\!\}\cdot[\![\boldsymbol{\sigma}_h]\!] - 2\int_{\mathcal{E}_h^i}\mathbf{S}\{\!\{\mathbf{u}_h\}\!\}\cdot\{\!\{\mathbf{v}_h\}\!\}.$$

In this way, the foregoing equations together with (2.9a), (2.9b), (2.9c), (2.9e) and (2.9f) lead to the problem: Find $((\mathbf{t}_h,\boldsymbol{\sigma}_h),\mathbf{u}_h)\in H_h\times V_h$ such that

$$\begin{aligned}[\mathcal{A}_h(\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{s}_h,\boldsymbol{\tau}_h)] + [\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{u}_h] &= [\mathcal{F}_h,(\mathbf{s}_h,\boldsymbol{\tau}_h)] \quad \forall\,(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h,\\ [\mathcal{B}_h(\mathbf{t}_h,\boldsymbol{\sigma}_h),\mathbf{v}_h] - [\mathcal{S}_h(\mathbf{u}_h),\mathbf{v}_h] &= [\mathcal{G}_h,\mathbf{v}_h] + [\mathcal{C}_h(\mathbf{u}_h),\mathbf{v}_h] \quad \forall\,\mathbf{v}_h\in V_h,\end{aligned} \tag{2.11}$$

where $H_h := S_h\times\Sigma_h$, and the operators $\mathcal{A}_h : H_h\to H_h'$, $\mathcal{B}_h : H_h\to V_h'$, $\mathcal{S}_h : V_h\to V_h'$ and $\mathcal{C}_h : V_h\to V_h'$, and the functionals $\mathcal{F}_h : H_h\to R$ and $\mathcal{G}_h : V_h\to R$, are defined by

$$\begin{aligned}[\mathcal{A}_h(\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{s}_h,\boldsymbol{\tau}_h)] :=\ & (\boldsymbol{\psi}(\mathbf{t}_h),\mathbf{s}_h)_{\mathcal{T}_h} - (\mathbf{s}_h,\boldsymbol{\sigma}_h^{\mathrm{d}})_{\mathcal{T}_h} + (\mathbf{t}_h,\boldsymbol{\tau}_h^{\mathrm{d}})_{\mathcal{T}_h}\\ &+\frac{1}{2}\int_{\mathcal{E}_h^i}\mathbf{S}^{-1}[\![\boldsymbol{\sigma}_h]\!]\cdot[\![\boldsymbol{\tau}_h]\!]\\ &+\kappa_1(\boldsymbol{\sigma}_h^{\mathrm{d}}-\boldsymbol{\psi}(\mathbf{t}_h),\boldsymbol{\tau}_h^{\mathrm{d}})_{\mathcal{T}_h}+\kappa_2(\mathbf{div}_h(\boldsymbol{\sigma}_h),\mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h},\end{aligned} \tag{2.12}$$

$$[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{v}_h] := (\mathbf{v}_h,\mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} - \int_{\mathcal{E}_h^i}\{\!\{\mathbf{v}_h\}\!\}\cdot[\![\boldsymbol{\tau}_h]\!], \tag{2.13}$$

$$[\mathcal{S}_h(\mathbf{u}_h),\mathbf{v}_h] := \langle \mathbf{S}\mathbf{u}_h,\mathbf{v}_h\rangle_{\partial\mathcal{T}_h}, \tag{2.14}$$

$$[\mathcal{C}_h(\mathbf{u}_h),\mathbf{v}_h] := -2\int_{\mathcal{E}_h^i}\mathbf{S}\{\!\{\mathbf{u}_h\}\!\}\cdot\{\!\{\mathbf{v}_h\}\!\},$$
$$[\mathcal{F}_h,(\mathbf{s}_h,\boldsymbol{\tau}_h)] := \langle\boldsymbol{\tau}_h\boldsymbol{\nu},\mathbf{g}\rangle_\Gamma - \kappa_2(\mathbf{f},\mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h},$$
$$[\mathcal{G}_h,\mathbf{v}_h] := -(\mathbf{f},\mathbf{v}_h)_{\mathcal{T}_h} - \langle\mathbf{S}\mathbf{g},\mathbf{v}_h\rangle_\Gamma,$$

where $[\cdot,\cdot]$ stands in each case for the duality pairing induced by the corresponding operators and functionals. Note, for purposes that will become clear below, that the expression $[\mathcal{C}_h(\mathbf{u}_h),\mathbf{v}_h]$ has been placed on the right-hand side of the second equation in (2.11). In addition, while the above operators and functionals are defined on discrete spaces, it is not difficult to see that they can act on continuous spaces as well. For example, $\mathcal{A}_h$ can actually be defined on $(S_h+\mathbb{L}^2(\Omega))\times(\Sigma_h+\mathbb{H}(\mathbf{div};\Omega))$ and similarly for the other ones. In particular, this fact will be employed at the beginning of Sect. 4.

3 Solvability Analysis

In this section, we establish the unique solvability of the nonlinear problem (2.11). To this end, and following [3,4], we let $\mathtt{h}\in L^\infty(\mathcal{E}_h)$ be the function related to the local meshsizes, that is

$$\mathtt{h}(x) := \begin{cases}\min\{h_{T_1},h_{T_2}\} & \text{if } x\in\mathrm{int}(\partial T_1\cap\partial T_2),\\ h_T & \text{if } x\in\mathrm{int}(\partial T\cap\Gamma),\end{cases}$$

and assume that the meshsize is bounded, that is, that there exists a constant $h_0>0$ such that

$$h := \max_{T\in\mathcal{T}_h}\{h_T\} \;\le\; h_0. \tag{3.1}$$

The main idea of our analysis consist of redefining (2.11) as a fixed point problem.

3.1 Preliminaries

The analysis below requires the following preliminary results.

Lemma 3.1 (Discrete trace's inequality) *There exists $C_{\rm tr} > 0$, depending only on the shape regularity of the mesh, such that for each $T \in \mathcal{T}_h$ and $F \in \partial T$ there holds*

$$\|\mathbf{z}\|^2_{0,F} \le C_{\rm tr}\left\{h_T^{-1}\|\mathbf{z}\|^2_{0,T} \;+\; h_T|\mathbf{z}|^2_{1,T}\right\} \quad \forall\,\mathbf{z}\in\mathbf{H}^1(T). \tag{3.2}$$

Proof The proof uses a trace theorem and a scaling argument (see [8] for details). □

Lemma 3.2 *There exists $c_0 > 0$, independent of h, such that for all $\mathbf{z} \in \mathbf{H}^1(\Omega)$ there holds*

$$\|\mathrm{h}^{1/2}\mathbf{z}\|_{0,\mathcal{E}_h^i} \le c_0\|\mathbf{z}\|_{1,\Omega}. \tag{3.3}$$

Proof Given $\mathbf{z}\in\mathbf{H}^1(\Omega)$, we have

$$\begin{aligned}\|\mathrm{h}^{1/2}\mathbf{z}\|^2_{0,\mathcal{E}_h^i} &= \int_{\mathcal{E}_h^i}\mathrm{h}|\mathbf{z}|^2 \;=\; \frac{1}{2}\int_{\mathcal{E}_h^i}\mathrm{h}\left(|\mathbf{z}^+|^2+|\mathbf{z}^-|^2\right) \;\le\; \frac{1}{2}\sum_{T\in\mathcal{T}_h}\int_{\partial T}\mathrm{h}|\mathbf{z}|^2\\ &\le C\sum_{T\in\mathcal{T}_h} h_T\|\mathbf{z}\|^2_{0,\partial T},\end{aligned}$$

where C depends on the regularity of $\mathcal{T}_h$. Next, using (3.2) and (3.1), we deduce from the foregoing inequalities that

$$\begin{aligned}\|\mathrm{h}^{1/2}\mathbf{z}\|^2_{0,\mathcal{E}_h^i} &\le CC_{\rm tr}\sum_{T\in\mathcal{T}_h} h_T\left\{h_T^{-1}\|\mathbf{z}\|^2_{0,T}+h_T|\mathbf{z}|^2_{1,T}\right\}\\ &\le CC_{\rm tr}(1+h^2)\sum_{T\in\mathcal{T}_h}\|\mathbf{z}\|^2_{1,T} \;\le\; c_0^2\|\mathbf{z}\|^2_{1,\Omega},\end{aligned}$$

with $c_0 := (CC_{\rm tr}(1+h_0^2))^{1/2}$, which completes the proof. □

Lemma 3.3 *There exists a constant $c_1 > 0$, independent of h, such that*

$$\|\boldsymbol{\tau}_h\|^2_{0,\Omega} \;\le\; c_1\left\{\|\boldsymbol{\tau}_h^{\rm d}\|^2_{0,\Omega} \;+\; \|\mathbf{div}_h(\boldsymbol{\tau}_h)\|^2_{\mathcal{T}_h} \;+\; \|\mathrm{h}^{-1/2}[\![\,\boldsymbol{\tau}_h\,]\!]\|^2_{0,\mathcal{E}_h^i}\right\} \quad \forall\,\boldsymbol{\tau}_h\in\Sigma_h.$$

Proof We follow similarly as in the proof of [2, Proposition 3.1, Chapter IV]. Indeed, given $\boldsymbol{\tau}_h\in\Sigma_h$, we know from [22, Corollary 2.4 in Chapter I] that there is a unique $\mathbf{z}\in\mathbf{H}^1_0(\Omega)$ such that $\mathrm{div}(\mathbf{z}) = \mathrm{tr}\,(\boldsymbol{\tau}_h)$ and

$$\|\mathbf{z}\|_{1,\Omega} \le C\|\mathrm{tr}\,(\boldsymbol{\tau}_h)\,\|_{0,\Omega}. \tag{3.4}$$

Now, utilizing that $\mathbf{z}\in\mathbf{H}^1_0(\Omega)$, we have that

$$\begin{aligned}\|\mathrm{tr}\,(\boldsymbol{\tau}_h)\,\|^2_{0,\Omega} &= \int_\Omega \mathrm{tr}\,(\boldsymbol{\tau}_h)\,\mathrm{div}(\mathbf{z}) \;=\; \int_\Omega \boldsymbol{\tau}_h : \{\mathrm{tr}\,(\nabla\mathbf{z})\,\mathbb{I}\}\,,\\ &= n\int_\Omega \boldsymbol{\tau}_h : (\nabla\mathbf{z}-(\nabla\mathbf{z})^{\rm d}) \;=\; n\int_\Omega \boldsymbol{\tau}_h : \nabla\mathbf{z} - n\int_\Omega \boldsymbol{\tau}_h^{\rm d} : \nabla\mathbf{z},\end{aligned}$$

$$= n \sum_{T \in \mathcal{T}_h} \left\{ - \int_T \mathbf{z} \cdot \mathbf{div}(\boldsymbol{\tau}_h) + \int_{\partial T} \boldsymbol{\tau}_h \boldsymbol{\nu} \cdot \mathbf{z} \right\} - n \int_\Omega \boldsymbol{\tau}_h^{\mathrm{d}} : \nabla \mathbf{z},$$

$$= -n(\mathbf{z}, \mathbf{div}_h(\boldsymbol{\tau}_h))_{\mathcal{T}_h} + n \int_{\mathcal{E}_h^i} [\![\boldsymbol{\tau}_h]\!] \cdot \mathbf{z} - n \int_\Omega \boldsymbol{\tau}_h^{\mathrm{d}} : \nabla \mathbf{z}.$$

Next, applying Cauchy–Schwarz inequality, and then (3.3) and (3.4), we find that

$$\begin{aligned}
\| \mathrm{tr}\,(\boldsymbol{\tau}_h) \|_{0,\Omega}^2 &\leq n \|\mathbf{z}\|_{0,\Omega} \|\mathbf{div}_h(\boldsymbol{\tau}_h)\|_{\mathcal{T}_h} + n \|\mathrm{h}^{-1/2} [\![\boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i} \|\mathrm{h}^{1/2} \mathbf{z}\|_{0,\mathcal{E}_h^i} + n \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega} |\mathbf{z}|_{1,\Omega} \\
&\leq n \|\mathbf{z}\|_{0,\Omega} \|\mathbf{div}_h(\boldsymbol{\tau}_h)\|_{\mathcal{T}_h} + n c_0 \|\mathrm{h}^{-1/2} [\![\boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i} \|\mathbf{z}\|_{1,\Omega} + n \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega} |\mathbf{z}|_{1,\Omega} \\
&\leq C \|\mathbf{z}\|_{1,\Omega} \left\{ \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}^2 + \|\mathbf{div}_h(\boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2 + \|\mathrm{h}^{-1/2} [\![\boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \right\}^{1/2} \\
&\leq C \| \mathrm{tr}\,(\boldsymbol{\tau}_h) \|_{0,\Omega} \left\{ \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}^2 + \|\mathbf{div}_h(\boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2 + \|\mathrm{h}^{-1/2} [\![\boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \right\}^{1/2},
\end{aligned}$$

which gives

$$\| \mathrm{tr}\,(\boldsymbol{\tau}_h) \|_{0,\Omega}^2 \leq C \left\{ \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}^2 + \|\mathbf{div}_h(\boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2 + \|\mathrm{h}^{-1/2} [\![\boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \right\}.$$

This inequality and the fact that $\|\boldsymbol{\tau}_h\|_{0,\Omega}^2 = \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}^2 + \frac{1}{n} \| \mathrm{tr}\,(\boldsymbol{\tau}_h) \|_{0,\Omega}^2$, complete the proof. □

We now realize, thanks to the previous lemma, that for convenience of further analysis, we need to establish a particular choice of the stabilization tensor $\mathbf{S}$. For this purpose, we let $\tau > 0$ be a constant and set the tensor $\mathbf{S}$ as follows

$$\mathbf{S}|_F := \tau \, \mathrm{h} \, \mathbb{I} \quad \forall \, F \in \mathcal{E}_h, \tag{3.5}$$

which certainly yields

$$\mathbf{S}^{-1}|_F := (\tau \mathrm{h})^{-1} \mathbb{I} \quad \forall \, F \in \mathcal{E}_h. \tag{3.6}$$

The parameter τ introduced here will play a key role later on for proving that the fixed point operator derived from our solvability analysis is in fact a contraction (see Lemma 3.12 below). In addition, we consider the following definition of a norm onto Σ_h

$$\|\boldsymbol{\tau}_h\|_{\Sigma_h}^2 := \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}^2 + \|\mathbf{div}_h(\boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2 + \|(\tau \mathrm{h})^{-1/2} [\![\boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \quad \forall \, \boldsymbol{\tau}_h \in \Sigma_h$$

which, according to Lemma 3.3, satisfies

$$\|\boldsymbol{\tau}_h\|_{0,\Omega} \leq c_2 \|\boldsymbol{\tau}_h\|_{\Sigma_h} \quad \forall \, \boldsymbol{\tau}_h \in \Sigma_h, \tag{3.7}$$

where $c_2^2 := c_1 \max\{1, \tau\} > 0$ is independent of h. Note that the above suggests the following norm on $H_h := S_h \times \Sigma_h$

$$\|(\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h} := \left\{ \|\mathbf{s}_h\|_{0,\Omega}^2 + \|\boldsymbol{\tau}_h\|_{\Sigma_h}^2 \right\}^{1/2} \quad \forall \, (\mathbf{s}_h, \boldsymbol{\tau}_h) \in H_h.$$

On the other hand, we define the nonlinear operator $\mathbb{A} : S_h \to S_h'$ by

$$[\mathbb{A}(\mathbf{t}_h), \mathbf{s}_h] := (\boldsymbol{\psi}(\mathbf{t}_h), \mathbf{s}_h)_{\mathcal{T}_h} \quad \forall \, \mathbf{t}_h, \mathbf{s}_h \in S_h.$$

Then, we have the following result.

Lemma 3.4 *Let γ_0 and α_0 be the constants from (2.2) and (2.3), respectively. Then, for all $\mathbf{t}_h, \mathbf{s}_h \in S_h$ there hold*

$$\|\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h)\|_{S_h'} \leq \gamma_0 \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega} \tag{3.8}$$

and

$$[\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h), \mathbf{t}_h - \mathbf{s}_h] \geq \alpha_0 \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega}^2. \tag{3.9}$$

Proof See [21, Lemma 2.1] or [4, Section 3]. □

We are now ready to establish that the nonlinear operator $\mathcal{A}_h$ defining the problem (2.11) is also Lipschitz-continuous and strongly monotone. In particular, the second property will depend on a suitable choice of the parameter κ_1.

Lemma 3.5 *Let $\mathcal{A}_h$ be the nonlinear operator defined by (2.12). Then, there exists a constant $C_{\rm LC} > 0$, independent of h and τ, such that*

$$\|\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h'} \leq C_{\rm LC} \|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h} \quad \forall\, (\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h) \in H_h.$$

Proof Given $(\mathbf{t}_h, \boldsymbol{\sigma}_h)$, $(\mathbf{s}_h, \boldsymbol{\tau}_h)$ and $(\mathbf{r}_h, \boldsymbol{\rho}_h) \in H_h$, we obtain from the definition of $\mathbb{A}$ and (3.6) that

$$\begin{aligned}
[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{r}_h, \boldsymbol{\rho}_h)] \;=\; & [\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h), \mathbf{r}_h] - \kappa_1 [\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h), \boldsymbol{\rho}_h^{\rm d}] \\
& - (\mathbf{r}_h, (\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d})_{\mathcal{T}_h} + (\mathbf{t}_h - \mathbf{s}_h, \boldsymbol{\rho}_h^{\rm d})_{\mathcal{T}_h} \\
& + \frac{1}{2} \sum_{F \in \mathcal{E}_h^i} \int_F (\tau \mathrm{h})^{-1/2} [\![(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)]\!] \cdot (\tau \mathrm{h})^{-1/2} [\![\boldsymbol{\rho}_h]\!] \\
& + \kappa_1 ((\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}, \boldsymbol{\rho}_h^{\rm d})_{\mathcal{T}_h} + \kappa_2 (\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h), \mathbf{div}_h(\boldsymbol{\rho}_h))_{\mathcal{T}_h},
\end{aligned} \tag{3.10}$$

from which, applying Cauchy–Schwarz inequality and (3.8), it follows that

$$\begin{aligned}
[\mathcal{A}(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}(\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{r}_h, \boldsymbol{\rho}_h)] \;\leq\; & \|\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h)\|_{S_h'} \|\mathbf{r}_h\|_{0,\Omega} \\
& + \kappa_1 \|\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h)\|_{S_h'} \|\boldsymbol{\rho}_h^{\rm d}\|_{0,\Omega} \\
& + \|\mathbf{r}_h\|_{0,\Omega} \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega} + \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega} \|\boldsymbol{\rho}_h^{\rm d}\|_{0,\Omega} \\
& + \frac{1}{2} \|(\tau \mathrm{h})^{-1/2} [\![(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)]\!]\|_{0,\mathcal{E}_h^i} \|(\tau \mathrm{h})^{-1/2} [\![\boldsymbol{\rho}_h]\!]\|_{0,\mathcal{E}_h^i} + \kappa_1 \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega} \|\boldsymbol{\rho}_h^{\rm d}\|_{0,\Omega} \\
& + \kappa_2 \|\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)\|_{\mathcal{T}_h} \|\mathbf{div}_h(\boldsymbol{\rho}_h)\|_{\mathcal{T}_h}, \\
\leq\; & \gamma_0 \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega} \|\mathbf{r}_h\|_{0,\Omega} + \gamma_0 \kappa_1 \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega} \|\boldsymbol{\rho}_h^{\rm d}\|_{0,\Omega} \\
& + \|\mathbf{r}_h\|_{0,\Omega} \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega} + \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega} \|\boldsymbol{\rho}_h^{\rm d}\|_{0,\Omega} \\
& + \frac{1}{2} \|(\tau \mathrm{h})^{-1/2} [\![(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)]\!]\|_{0,\mathcal{E}_h^i} \|(\tau \mathrm{h})^{-1/2} [\![\boldsymbol{\rho}_h]\!]\|_{0,\mathcal{E}_h^i} + \kappa_1 \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega} \|\boldsymbol{\rho}_h^{\rm d}\|_{0,\Omega} \\
& + \kappa_2 \|\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)\|_{\mathcal{T}_h} \|\mathbf{div}_h(\boldsymbol{\rho}_h)\|_{\mathcal{T}_h}.
\end{aligned}$$

In this way, setting

$$C_{\rm LC} \;:=\; 3 \max\{1, \gamma_0, \kappa_1, \gamma_0 \kappa_1, \kappa_2\},$$

we conclude that

$$[\mathcal{A}(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}(\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{r}_h, \boldsymbol{\rho}_h)] \leq C_{\rm LC} \|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h} \|(\mathbf{r}_h, \boldsymbol{\rho}_h)\|_{H_h},$$

which ends the proof. □

Lemma 3.6 *Let $\mathcal{A}_h$ be the nonlinear operator defined by (2.12), and assume that the parameter κ_1 lies in $\left(0, \frac{2\alpha_0}{\gamma_0^2}\right)$, where α_0 and γ_0 are the positive constants from (2.2) and (2.3). Then, there exists a constant $C_{\rm SM} > 0$, independent of h and τ, such that*

$$[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)] \;\geq\; C_{\rm SM}\, \|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h}^2,$$

for all $(\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h) \in H_h$.

Proof Given $(\mathbf{t}_h, \boldsymbol{\sigma}_h)$ and $(\mathbf{s}_h, \boldsymbol{\tau}_h) \in H_h$, we take $(\mathbf{r}_h, \boldsymbol{\rho}_h) = (\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)$ in (3.10), to obtain

$$\begin{aligned}
[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)] \;&=\; [\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h), \mathbf{t}_h - \mathbf{s}_h] \\
&\quad - \kappa_1[\mathbb{A}(\mathbf{t}_h) - \mathbb{A}(\mathbf{s}_h), (\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}] + \frac{1}{2}\|(\tau \mathrm{h})^{-1/2}[\![\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \\
&\quad + \kappa_1\|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega}^2 + \kappa_2\|\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2,
\end{aligned}$$

which, according to (3.8) and (3.9), implies that

$$\begin{aligned}
&[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)] \\
&\geq \alpha_0\, \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega}^2 - \gamma_0\kappa_1\, \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega}\, \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega} + \frac{1}{2}\|(\tau \mathrm{h})^{-1/2}[\![\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \\
&\quad + \kappa_1\, \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega}^2 + \kappa_2\, \|\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2, \\
&\geq \alpha_0\, \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega}^2 - \gamma_0\kappa_1 \left\{ \frac{\|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega}^2}{2\delta} + \frac{\delta\, \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega}^2}{2} \right\} \\
&\quad + \frac{1}{2}\|(\tau \mathrm{h})^{-1/2}[\![\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 + \kappa_1\, \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega}^2 + \kappa_2\, \|\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2, \\
&= \left(\alpha_0 - \frac{\gamma_0\kappa_1}{2\delta}\right) \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega}^2 + \kappa_1 \left(1 - \frac{\gamma_0\delta}{2}\right) \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega}^2 \\
&\quad + \kappa_2\, \|\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2 + \frac{1}{2}\|(\tau \mathrm{h})^{-1/2}[\![\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \quad \forall\, \delta > 0.
\end{aligned}$$

It follows that the constants multiplying the norms above become positive if $\delta \in \left(0, \frac{2}{\gamma_0}\right)$ and $\kappa_1 \in \left(0, \frac{2\alpha_0\delta}{\gamma_0}\right)$. In particular, for $\delta = \frac{1}{\gamma_0}$ we require $\kappa_1 \in \left(0, \frac{2\alpha_0}{\gamma_0^2}\right)$, whence we find that

$$\begin{aligned}
&[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)] \\
&\geq \left(\alpha_0 - \frac{\gamma_0^2\kappa_1}{2}\right) \|\mathbf{t}_h - \mathbf{s}_h\|_{0,\Omega}^2 + \frac{\kappa_1}{2}\|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)^{\rm d}\|_{0,\Omega}^2 \\
&\quad + \kappa_2\, \|\mathbf{div}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h)\|_{\mathcal{T}_h}^2 + \frac{1}{2}\|(\tau \mathrm{h})^{-1/2}[\![\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h]\!]\|_{0,\mathcal{E}_h^i}^2 \\
&\geq C_{\rm SM}\|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h}^2,
\end{aligned}$$

with $C_{\rm SM} := \min\left\{\alpha_0 - \frac{\gamma_0^2\kappa_1}{2}, \frac{\kappa_1}{2}, \kappa_2, \frac{1}{2}\right\}$, thus completing the proof of the lemma. □

Our next goal is to show the discrete inf-sup condition for the linear operator $\mathcal{B}_h$. More precisely, we have the following result.

Lemma 3.7 *There exists a constant $C_{\rm inf} > 0$, independent of h and τ, such that*

$$\sup_{\substack{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h\\(\mathbf{s}_h,\boldsymbol{\tau}_h)\neq\mathbf{0}}} \frac{[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{v}_h]}{\|(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}} \geq C_{\rm inf}\,\|\mathbf{v}_h\|_{0,\Omega} \quad \forall\,\mathbf{v}_h\in V_h.$$

Proof We begin by recalling from (2.13) that $\mathcal{B}_h$ does not depend on $\mathbf{s}_h$, and hence it suffices to show the existence of $C_{\rm inf} > 0$ such that

$$\sup_{\substack{\boldsymbol{\tau}_h\in\Sigma_h\\\boldsymbol{\tau}_h\neq\mathbf{0}}} \frac{\displaystyle\int_\Omega \mathbf{v}_h\cdot\mathbf{div}_h(\boldsymbol{\tau}_h) - \int_{\mathcal{E}_h^i}\{\!\!\{\mathbf{v}_h\}\!\!\}\cdot[\![\boldsymbol{\tau}_h]\!]}{\|\boldsymbol{\tau}_h\|_{\Sigma_h}} \geq C_{\rm inf}\,\|\mathbf{v}_h\|_{0,\Omega} \quad \forall\,\mathbf{v}_h\in V_h.$$

To this end we let $\mathrm{RT}_{k-1}(\Omega)$ be the global Raviart–Thomas space of degree $k-1$, which is clearly contained in S_h, and note that

$$\sup_{\substack{\boldsymbol{\tau}_h\in\Sigma_h\\\boldsymbol{\tau}_h\neq\mathbf{0}}} \frac{\displaystyle\int_\Omega \mathbf{v}_h\cdot\mathbf{div}_h(\boldsymbol{\tau}_h) - \int_{\mathcal{E}_h^i}\{\!\!\{\mathbf{v}_h\}\!\!\}\cdot[\![\boldsymbol{\tau}_h]\!]}{\|\boldsymbol{\tau}_h\|_{\Sigma_h}} \geq \sup_{\substack{\boldsymbol{\tau}_h\in\mathrm{RT}_{k-1}(\Omega)\setminus\{\mathbf{0}\}\\\int_\Omega \mathrm{tr}(\boldsymbol{\tau}_h)=0}} \frac{\displaystyle\int_\Omega \mathbf{v}_h\cdot\mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\Sigma_h}}.$$

In this way, and observing that $\|\boldsymbol{\tau}_h\|_{\Sigma_h}$ is equivalent to $\|\boldsymbol{\tau}_h\|_{\mathbf{div},\Omega}$ $\forall\,\boldsymbol{\tau}_h\in\mathrm{RT}_{k-1}(\Omega)$ such that $\int_\Omega \mathrm{tr}\,(\boldsymbol{\tau}_h)=0$, with constants independent of h and τ, the rest of the proof follows from classical results from mixed finite element methods (see, e.g. [18, Section 4.2 and Lemma 2.6]). □

The following three lemmas establish the positive semidefiniteness of $\mathcal{S}_h$ and some discrete trace, inverse, and boundedness inequalities to be employed later on.

Lemma 3.8 *The operator $\mathcal{S}_h : V_h \to V_h'$, defined by (2.14) is positive semidefinite, that is,*

$$[\mathcal{S}_h(\mathbf{v}_h),\mathbf{v}_h] \geq 0 \quad \forall\,\mathbf{v}_h\in V_h.$$

Proof It is clear from (2.14) that

$$[\mathcal{S}_h(\mathbf{v}_h),\mathbf{v}_h] = \sum_{T\in\mathcal{T}_h}\sum_{F\in\partial T}\int_F \mathbf{S}\mathbf{v}_h\cdot\mathbf{v}_h \quad \forall\,\mathbf{v}_h\in V_h,$$

which, thanks to the fact that $\mathbf{S}$ is a positive definite tensor on $\mathcal{E}_h$, completes the proof. □

Lemma 3.9 (Discrete trace's inequality + inverse's inequality) *There exists $C_{\rm inv} > 0$, depending only on k and the shape regularity of the mesh, such that*

$$\|\mathbf{v}\|_{0,\partial T}^2 \leq C_{\rm inv}\,h_T^{-1}\|\mathbf{v}\|_{0,T}^2 \quad \forall\,\mathbf{v}\in\mathbf{P}_k(T), \quad \forall\,T\in\mathcal{T}_h, \tag{3.11}$$

and

$$\|\boldsymbol{\tau}\|_{0,\partial T}^2 \leq C_{\rm inv}\,h_T^{-1}\|\boldsymbol{\tau}\|_{0,T}^2 \quad \forall\,\boldsymbol{\tau}\in\mathbb{P}_k(T), \quad \forall\,T\in\mathcal{T}_h. \tag{3.12}$$

Proof The proof uses the discrete trace inequality from Lemma 3.1 and an inverse inequality. See also [3, Lemma 3.2]. □

Lemma 3.10 *There exist constants $\widehat{C}_1, \widehat{C}_2, \widehat{C}_3 > 0$, independent of h and τ, such that*

(i) $\|\mathrm{h}^{1/2}\{\!\!\{\mathbf{v}_h\}\!\!\}\|_{0,\mathcal{E}_h^i} \leq \widehat{C}_1\|\mathbf{v}_h\|_{0,\Omega}$ $\forall\,\mathbf{v}_h\in V_h$.

(ii) $\|\mathrm{h}^{1/2}\mathbf{v}_h\|_{0,\mathcal{E}_h^\partial} \leq \widehat{C}_2\|\mathbf{v}_h\|_{0,\Omega} \quad \forall\, \mathbf{v}_h \in V_h.$

(iii) $\|\mathrm{h}^{1/2}\boldsymbol{\tau}_h\boldsymbol{\nu}\|_{0,\mathcal{E}_h^\partial} \leq \widehat{C}_3\|\boldsymbol{\tau}_h\|_{0,\Omega} \quad \forall\, \boldsymbol{\tau}_h \in \Sigma_h.$

Proof Given $\mathbf{v}_h \in V_h$, we use (3.11) to deduce that

$$\|\mathrm{h}^{1/2}\{\mathbf{v}_h\}\|^2_{0,\mathcal{E}_h^i} = \frac{1}{4}\int_{\mathcal{E}_h^i} \mathrm{h}|\mathbf{v}_h^+ + \mathbf{v}_h^-|^2 \leq \frac{1}{2}\int_{\mathcal{E}_h^i} \mathrm{h}\left(|\mathbf{v}_h^+|^2 + |\mathbf{v}_h^-|^2\right) \leq \frac{1}{2}\sum_{T\in\mathcal{T}_h}\int_{\partial T} \mathrm{h}|\mathbf{v}_h|^2$$
$$\leq C_1 \sum_{T\in\mathcal{T}_h} h_T\|\mathbf{v}_h\|^2_{0,\partial T} \leq C_1 C_{\mathrm{inv}} \sum_{T\in\mathcal{T}_h} \|\mathbf{v}_h\|^2_{0,T} = C_1 C_{\mathrm{inv}}\|\mathbf{v}_h\|^2_{0,\Omega},$$

which shows (i) with $\widehat{C}_1 := (C_1 C_{\mathrm{inv}})^{1/2} > 0$. Next, using that $\mathrm{h} = h_T$ on $\mathcal{E}_h^\partial$, and applying again (3.11), we find that

$$\|\mathrm{h}^{1/2}\mathbf{v}_h\|^2_{0,\mathcal{E}_h^\partial} = \int_{\mathcal{E}_h^\partial} \mathrm{h}|\mathbf{v}_h|^2 \leq \sum_{T\in\mathcal{T}_h} h_T\|\mathbf{v}_h\|^2_{0,\partial T} \leq C_{\mathrm{inv}}\|\mathbf{v}_h\|^2_{0,\Omega},$$

which proves (ii) with $\widehat{C}_2 := (C_{\mathrm{inv}})^{1/2}$. Finally, the proof of (iii) follows from (3.12). □

Using Lemma 3.10, the definition of tensor $\mathbf{S}$ given in (3.5), and the Cauchy–Schwarz inequality, it is easy to check that the operators $\mathcal{B}_h$, $\mathcal{S}_h$ and $\mathcal{C}_h$, and the functionals $\mathcal{F}_h$ and $\mathcal{G}_h$, are all bounded with respect to the corresponding norms. More precisely, the corresponding bounds are established in the following lemma.

Lemma 3.11 *Let $\mathbf{s}_h \in S_h$, $\boldsymbol{\tau}_h \in \Sigma_h$ and $\mathbf{u}_h, \mathbf{v}_h \in V_h$. Then there hold*

$$\begin{aligned}
|[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{v}_h]| &\leq (1+\tau\widehat{C}_1)\ \|(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}\ \|\mathbf{v}_h\|_{0,\Omega}\\
|[\mathcal{S}_h(\mathbf{u}_h),\mathbf{v}_h]| &\leq \tau\widehat{C}_1\ \|\mathbf{u}_h\|_{0,\Omega}\ \|\mathbf{v}_h\|_{0,\Omega}\\
|[\mathcal{C}_h(\mathbf{u}_h),\mathbf{v}_h]| &\leq 2\tau\widehat{C}_1^2\ \|\mathbf{u}_h\|_{0,\Omega}\ \|\mathbf{v}_h\|_{0,\Omega}\\
|[\mathcal{F}_h,(\mathbf{s}_h,\boldsymbol{\tau}_h)]| &\leq \left(\kappa_2 + c_2\widehat{C}_3\right)\mathbb{B}(\mathbf{f},\mathbf{g})\ \|(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}\\
|[\mathcal{G}_h,\mathbf{v}_h]| &\leq \left(1+\tau h_0\widehat{C}_2\right)\mathbb{B}(\mathbf{f},\mathbf{g})\ \|\mathbf{v}_h\|_{0,\Omega}
\end{aligned} \tag{3.13}$$

where

$$\mathbb{B}(\mathbf{f},\mathbf{g}) := \|\mathbf{f}\|_{0,\Omega} + \|\mathrm{h}^{-1/2}\mathbf{g}\|_{0,\mathcal{E}_h^\partial}.$$

Proof The proof uses Lemma 3.10 and the definitions of each operator and functional. We omit further details and refer to [3, Lemma 4.4]. □

We end this section, by recalling from [20] the following abstract theorem.

Theorem 3.1 *Let X, M be Hilbert spaces and assume that*

(i) *the operator $\mathcal{A} : X \to X'$ is Lipschitz continuous and strongly monotonic, that is, there exist $\gamma, \alpha > 0$ such that*

$$\|\mathcal{A}(\mathbf{s}_1) - \mathcal{A}(\mathbf{s}_2)\|_{X'} \leq \gamma\ \|\mathbf{s}_1 - \mathbf{s}_2\|_X \quad \forall\, \mathbf{s}_1, \mathbf{s}_2 \in X$$

and

$$[\mathcal{A}(\mathbf{s}_1) - \mathcal{A}(\mathbf{s}_2), \mathbf{s}_1 - \mathbf{s}_2] \geq \alpha\ \|\mathbf{s}_1 - \mathbf{s}_2\|^2_X \quad \forall\, \mathbf{s}_1, \mathbf{s}_2 \in X;$$

(ii) *the linear operator $\mathcal{S}$ is positive semidefinite on M, that is*

$$[\mathcal{S}(\boldsymbol{\tau}), \boldsymbol{\tau}] \geq 0 \quad \forall\, \boldsymbol{\tau} \in M;$$

(iii) *the linear operator $\mathcal{B}$ satisfies an inf-sup condition on $X \times M$, that is, there exists $\beta > 0$ such that*

$$\sup_{\substack{\mathbf{s}\in X\\ \mathbf{s}\neq\mathbf{0}}} \frac{[\mathcal{B}(\mathbf{s}),\boldsymbol{\tau}]}{\|\mathbf{s}\|_X} \;\geq\; \beta\,\|\boldsymbol{\tau}\|_M \quad \forall\,\boldsymbol{\tau}\in M.$$

Then, given $\mathcal{F}\in X'$ and $\mathcal{G}\in M'$, there exists a unique solution $(\mathbf{t},\boldsymbol{\sigma})\in X\times M$ of

$$\begin{aligned}[\mathcal{A}(\mathbf{t}),\mathbf{s}] + [\mathcal{B}^*(\boldsymbol{\sigma}),\mathbf{s}] &= [\mathcal{F},\mathbf{s}] \quad \forall\,\mathbf{s}\in X,\\ [\mathcal{B}(\mathbf{t}),\boldsymbol{\tau}] - [\mathcal{S}(\boldsymbol{\sigma}),\boldsymbol{\tau}] &= [\mathcal{G},\boldsymbol{\tau}] \quad \forall\,\boldsymbol{\tau}\in M.\end{aligned}$$

In addition, the following estimates hold

$$\begin{aligned}\|\mathbf{t}\|_X &\leq C_1\left\{\|\mathcal{F}\|_{X'} \;+\; \|\mathcal{G}\|_{M'} \;+\; \|\mathcal{A}(\mathbf{0})\|_{X'}\right\},\\ \|\boldsymbol{\sigma}\|_M &\leq C_2\left\{\|\mathcal{F}\|_{X'} \;+\; \|\mathcal{G}\|_{M'} \;+\; \|\mathcal{A}(\mathbf{0})\|_{X'}\right\},\end{aligned}$$

where

$$C_1 \;:=\; \frac{1}{\alpha} + \frac{\|\mathcal{B}\|}{\alpha}C_2 \quad \textit{and} \quad C_2 \;:=\; \frac{\gamma^2}{\alpha\beta^2}\left(1+\frac{\|\mathcal{B}\|}{\alpha}\right).$$

Proof See [20, Lemma 2.1], where it is easy to show the last estimates from expressions (2.8) and (2.9) in [20]. □

3.2 Main Result

In order to prove existence and uniqueness of solution of (2.11), we now introduce the nonlinear mapping $\mathbb{T}_h : H_h\times V_h \to H_h\times V_h$ that, given $((\mathbf{r}_h,\boldsymbol{\rho}_h),\mathbf{w}_h)\in H_h\times V_h$, defines $\mathbb{T}_h((\mathbf{r}_h,\boldsymbol{\rho}_h),\mathbf{w}_h) := ((\mathbf{t}_h,\boldsymbol{\sigma}_h),\mathbf{u}_h)\in H_h\times V_h$ as the unique solution of the problem

$$\begin{aligned}[\mathcal{A}_h(\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{s}_h,\boldsymbol{\tau}_h)] \;+\; [\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{u}_h] &= [\mathcal{F}_h,(\mathbf{s}_h,\boldsymbol{\tau}_h)] \quad \forall\,(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h,\\ [\mathcal{B}_h(\mathbf{t}_h,\boldsymbol{\sigma}_h),\mathbf{v}_h] \;-\; [\mathcal{S}_h(\mathbf{u}_h),\mathbf{v}_h] &= [\mathcal{G}_h,\mathbf{v}_h] \;+\; [\mathcal{C}_h(\mathbf{w}_h),\mathbf{v}_h] \quad \forall\,\mathbf{v}_h\in V_h.\end{aligned}$$

Note that actually $\mathbb{T}_h((\mathbf{r}_h,\boldsymbol{\rho}_h),\mathbf{w}_h)$ depends only on the third component $\mathbf{w}_h\in V_h$. In addition, bearing in mind Lemmas 3.5, 3.6, 3.7 and 3.8, it follows from Theorem 3.1 that $\mathbb{T}_h$ is well-defined and there holds

$$\|(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h} \leq \widehat{C}_a\widetilde{C}\;\mathbb{B}(\mathbf{f},\mathbf{g}) \;+\; 2\widehat{C}_1^2\widehat{C}_a\tau\;\|\mathbf{w}_h\|_{0,\Omega}, \tag{3.14}$$

and

$$\|\mathbf{u}_h\|_{0,\Omega} \leq \widehat{C}_b\widetilde{C}\;\mathbb{B}(\mathbf{f},\mathbf{g}) \;+\; 2\widehat{C}_1^2\widehat{C}_b\tau\;\|\mathbf{w}_h\|_{0,\Omega}, \tag{3.15}$$

where

$$\begin{aligned}\widetilde{C} &:= 1+\kappa_2+\tau h_0\widehat{C}_2 + c_1^{1/2}\widehat{C}_3(1+\tau)^{1/2},\\ \widehat{C}_a &:= \frac{1}{C_{\mathrm{SM}}}\left(1+(1+\tau\widehat{C}_1)\widehat{C}_b\right),\\ \widehat{C}_b &:= \frac{C_{\mathrm{LC}}^2}{C_{\mathrm{SM}}C_{\mathrm{inf}}^2}\left(1+\frac{1+\tau\widehat{C}_1}{C_{\mathrm{SM}}}\right),\end{aligned}$$

and the constants $\widehat{C}_1$, $\widehat{C}_2$, and $\widehat{C}_3$ are those from Lemma 3.10. Observe here that the identity $\mathcal{A}_h(\mathbf{0},\mathbf{0}) = (\mathbf{0},\mathbf{0})$ and Lemma 3.11 have been employed to establish the estimates (3.14) and

(3.15). Also, we remark that the relevance of the introduction of $\mathbb{T}_h$ has to do with the fact that any eventual solution of (2.11) becomes a fixed point of $\mathbb{T}_h$ and conversely. Moreover, the following lemma establishes that $\mathbb{T}_h$ is indeed a contraction mapping and hence, thanks to the Banach Fixed-Point Theorem, it has a unique fixed point in $H_h \times V_h$.

Lemma 3.12 *Assume that the parameter τ lies in $\left(0, \frac{1}{\theta}\right)$, where*

$$\theta := \left(\frac{2\widehat{C}_1^2}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(1 + \frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right) > 0.$$

Then, $\mathbb{T}_h$ is a contraction.

Proof Given $((\mathbf{t}_h, \boldsymbol{\sigma}_h), \mathbf{u}_h)$, $((\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h), \widetilde{\mathbf{u}}_h)$, $((\mathbf{r}_h, \boldsymbol{\rho}_h), \mathbf{w}_h)$, and $((\widetilde{\mathbf{r}}_h, \widetilde{\boldsymbol{\rho}}_h), \widetilde{\mathbf{w}}_h)$ in $H_h \times V_h$ such that

$$\mathbb{T}_h((\mathbf{r}_h, \boldsymbol{\rho}_h), \mathbf{w}_h) = ((\mathbf{t}_h, \boldsymbol{\sigma}_h), \mathbf{u}_h) \quad \text{and} \quad \mathbb{T}_h((\widetilde{\mathbf{r}}_h, \widetilde{\boldsymbol{\rho}}_h), \widetilde{\mathbf{w}}_h) = ((\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h), \widetilde{\mathbf{u}}_h),$$

we know from the definition of $\mathbb{T}_h$ that

$$[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h)] + [\mathcal{B}_h(\mathbf{s}_h, \boldsymbol{\tau}_h), \mathbf{u}_h - \widetilde{\mathbf{u}}_h] = 0, \tag{3.16a}$$

$$[\mathcal{B}_h(\mathbf{t}_h - \widetilde{\mathbf{t}}_h, \boldsymbol{\sigma}_h - \widetilde{\boldsymbol{\sigma}}_h), \mathbf{v}_h] - [\mathcal{S}_h(\mathbf{u}_h - \widetilde{\mathbf{u}}_h), \mathbf{v}_h] = [\mathcal{C}_h(\mathbf{w}_h - \widetilde{\mathbf{w}}_h), \mathbf{v}_h], \tag{3.16b}$$

for all $((\mathbf{s}_h, \boldsymbol{\tau}_h), \mathbf{v}_h) \in H_h \times V_h$. Next, taking $(\mathbf{s}_h, \boldsymbol{\tau}_h) = (\mathbf{t}_h - \widetilde{\mathbf{t}}_h, \boldsymbol{\sigma}_h - \widetilde{\boldsymbol{\sigma}}_h)$ and $\mathbf{v}_h = \mathbf{u}_h - \widetilde{\mathbf{u}}_h$, we obtain from (3.16) that

$$\begin{aligned}[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h), (\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h)] \\ + [\mathcal{S}_h(\mathbf{u}_h - \widetilde{\mathbf{u}}_h), \mathbf{u}_h - \widetilde{\mathbf{u}}_h] = -[\mathcal{C}_h(\mathbf{w}_h - \widetilde{\mathbf{w}}_h), \mathbf{u}_h - \widetilde{\mathbf{u}}_h].\end{aligned} \tag{3.17}$$

Then, using the strong monotonicity of $\mathcal{A}_h$, the fact that $\mathcal{S}_h$ is positive semidefinite, and the boundedness of $\mathcal{C}_h$ [cf. (3.13)], we deduce from (3.17) that

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h)\|_{H_h}^2 \leq \frac{2\tau\widehat{C}_1^2}{C_{\mathrm{SM}}} \|\mathbf{w}_h - \widetilde{\mathbf{w}}_h\|_{0,\Omega} \|\mathbf{u}_h - \widetilde{\mathbf{u}}_h\|_{0,\Omega}. \tag{3.18}$$

On the other hand, employing the inf-sup condition for $\mathcal{B}_h$ (cf. Lemma 3.7), (3.16a), and the Lipschitz-continuity of $\mathcal{A}_h$ (cf. Lemma 3.6), we find that

$$\begin{aligned}\|\mathbf{u}_h - \widetilde{\mathbf{u}}_h\|_{0,\Omega} &\leq \frac{1}{C_{\mathrm{inf}}} \sup_{\substack{(\mathbf{s}_h, \boldsymbol{\tau}_h) \in H_h \\ (\mathbf{s}_h, \boldsymbol{\tau}_h) \neq \mathbf{0}}} \frac{|[\mathcal{B}_h(\mathbf{s}_h, \boldsymbol{\tau}_h), \mathbf{u}_h - \widetilde{\mathbf{u}}_h]|}{\|(\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h}}, \\ &= \frac{1}{C_{\mathrm{inf}}} \sup_{\substack{(\mathbf{s}_h, \boldsymbol{\tau}_h) \in H_h \\ (\mathbf{s}_h, \boldsymbol{\tau}_h) \neq \mathbf{0}}} \frac{|-[\mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h) - \mathcal{A}_h(\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h)]|}{\|(\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h}}, \\ &\leq \frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}} \|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h)\|_{H_h},\end{aligned}$$

which, together with (3.18), implies that

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\widetilde{\mathbf{t}}_h, \widetilde{\boldsymbol{\sigma}}_h)\|_{H_h} \leq \left(\frac{2\tau\widehat{C}_1^2}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right) \|\mathbf{w}_h - \widetilde{\mathbf{w}}_h\|_{0,\Omega}$$

and

$$\|\mathbf{u}_h - \widetilde{\mathbf{u}}_h\|_{0,\Omega} \leq \left(\frac{2\tau\widehat{C}_1^2}{C_{\rm SM}}\right)\left(\frac{C_{\rm LC}}{C_{\rm inf}}\right)^2 \|\mathbf{w}_h - \widetilde{\mathbf{w}}_h\|_{0,\Omega}.$$

In this way, we conclude that

$$\|\mathbb{T}_h((\mathbf{r}_h,\boldsymbol{\rho}_h),\mathbf{w}_h) - \mathbb{T}_h((\widetilde{\mathbf{r}}_h,\widetilde{\boldsymbol{\rho}}_h),\widetilde{\mathbf{w}}_h)\|_{H_h\times V_h} \leq L\ \|((\mathbf{r}_h,\boldsymbol{\rho}_h),\mathbf{w}_h) - ((\widetilde{\mathbf{r}}_h,\widetilde{\boldsymbol{\rho}}_h),\widetilde{\mathbf{w}}_h)\|_{H_h\times V_h},$$

with $L := \tau\theta$. Finally, since $C_{\rm inf}$, $C_{\rm LC}$, and $C_{\rm SM}$, are independent of $\tau > 0$, we can choose $\tau \in \left(0, \frac{1}{\theta}\right)$, which insures that $\mathbb{T}_h$ is a contraction and completes the proof. □

Now we are ready to establish the main result of this section.

Theorem 3.2 *Assume that*

$$0 < \tau < \min\left\{\frac{1}{\theta}, \frac{1}{2}\left(\frac{C_{\rm SM}}{(1+C_{\rm SM})\theta + \widehat{C}_1}\right)\right\}.$$

Then, there exists a unique $((\mathbf{t}_h,\boldsymbol{\sigma}_h),\mathbf{u}_h) \in H_h \times V_h$ *solution of* (2.11). *Moreover, there holds*

$$\|(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{\Sigma_h} \leq C_{\rm a}\ \mathbb{B}(\mathbf{f},\mathbf{g}) \quad \textit{and} \quad \|\mathbf{u}_h\|_{0,\Omega} \leq C_{\rm b}\ \mathbb{B}(\mathbf{f},\mathbf{g}),$$

where

$$C_{\rm a} := \widehat{C}_a\ (\widetilde{C} + 2\widehat{C}_1^2 C_b \tau) \quad \textit{and} \quad C_{\rm b} := 2\widehat{C}_b\widetilde{C}.$$

Proof The unique solvability of (2.11) follows straightforwardly from its equivalence with the fixed-point equation for $\mathbb{T}_h$, the corresponding Banach Theorem, and the fact that $\mathbb{T}_h$ becomes a contraction when $\tau < \frac{1}{\theta}$ (cf. Lemma 3.12). Then, denoting by $((\mathbf{t}_h,\boldsymbol{\sigma}_h),\mathbf{u}_h) \in H_h \times V_h$ the unique solution of (2.11), we have from (3.14) and (3.15) that

$$\|(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h} \leq \widehat{C}_a\widetilde{C}\ \mathbb{B}(\mathbf{f},\mathbf{g}) + 2\widehat{C}_1^2\widehat{C}_a\tau\ \|\mathbf{u}_h\|_{0,\Omega} \tag{3.19}$$

and

$$\|\mathbf{u}_h\|_{0,\Omega} \leq \widehat{C}_b\widetilde{C}\ \mathbb{B}(\mathbf{f},\mathbf{g}) + 2\widehat{C}_1^2\widehat{C}_b\tau\ \|\mathbf{u}_h\|_{0,\Omega}. \tag{3.20}$$

It remain to handle the second term on the right-hand side of (3.20). For this purpose, we now note that

$$\begin{aligned}
2\widehat{C}_1^2\widehat{C}_b\tau &= 2\widehat{C}_1^2\frac{C_{\rm LC}^2}{C_{\rm SM}C_{\rm inf}^2}\left(1 + \frac{1+\tau\widehat{C}_1}{C_{\rm SM}}\right)\tau \\
&= \left(\frac{2\widehat{C}_1^2}{C_{\rm SM}}\right)\left(\frac{C_{\rm LC}}{C_{\rm inf}}\right)\left(\frac{C_{\rm LC}}{C_{\rm inf}}\right)\left(1 + \frac{1+\tau\widehat{C}_1}{C_{\rm SM}}\right)\tau \\
&\leq \theta\left(1 + \frac{1+\tau\widehat{C}_1}{C_{\rm SM}}\right)\tau = \left(\theta + \frac{\theta + (\theta\tau)\widehat{C}_1}{C_{\rm SM}}\right)\tau
\end{aligned}$$

which, using the assumption on τ, gives

$$2\widehat{C}_1^2\widehat{C}_b\tau < \left(\theta + \frac{\theta+\widehat{C}_1}{C_{\rm SM}}\right)\tau = \left(\frac{(1+C_{\rm SM})\theta + \widehat{C}_1}{C_{\rm SM}}\right)\tau < \frac{1}{2}.$$

In this way, replacing the foregoing inequality back into (3.20), we deduce that

$$\|\mathbf{u}_h\|_{0,\Omega} \leq 2\widehat{C}_b\widetilde{C}\ \mathbb{B}(\mathbf{f},\mathbf{g}) = C_{\rm b}\ \mathbb{B}(\mathbf{f},\mathbf{g}),$$

which, together with (3.19), yields

$$\|(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h} \leq \left(\widehat{C}_a\widetilde{C} + 2\widehat{C}_1^2\widehat{C}_a C_{\mathrm{b}}\tau\right)\mathbb{B}(\mathbf{f},\mathbf{g}) \;=\; C_{\mathrm{a}}\;\mathbb{B}(\mathbf{f},\mathbf{g}),$$

thus completing the proof of the theorem. □

4 A-Priori Error Analysis

We now aim to derive the a priori error estimates for the augmented HDG scheme (2.11). We begin by remarking that the eventual extension to the present nonlinear case of the projection-based error analysis developed in [11] (see also [14]) does not seem straightforward, precisely because of the nonlinearity, and hence in what follows we adopt a more classical approach. Next, since $\mathbf{u}\in\mathbf{L}^2(\Omega)$ and $\nabla\mathbf{u}=\mathbf{t}\in\mathbb{L}^2(\Omega)$ [cf. (2.4)], we observe that actually $\mathbf{u}\in\mathbf{H}^1(\Omega)$, which guarantees that the jump $[\![\mathbf{u}]\!]$ vanish on any interior face of $\mathcal{T}_h$ and there holds $\{\mathbf{u}\}=\mathbf{u}$. In addition, since $\boldsymbol{\sigma}=\boldsymbol{\psi}(\nabla\mathbf{u})-p\mathbb{I}\in\mathbb{L}^2(\Omega)$ and $\mathbf{div}(\boldsymbol{\sigma})=-\mathbf{f}$ in Ω, with $\mathbf{f}\in\mathbf{L}^2(\Omega)$, we conclude that $\boldsymbol{\sigma}\in\mathbb{H}(\mathbf{div};\Omega)$, whence $[\![\boldsymbol{\sigma}]\!]=\mathbf{0}$ on each $F\in\mathcal{E}_h^i$. Then, it is easy to check that $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u})$ satisfies the equations of (2.11), and then we obtain the error equations

$$[\mathcal{A}_h(\mathbf{t},\boldsymbol{\sigma})-\mathcal{A}_h(\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{s}_h,\boldsymbol{\tau}_h)]+[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{u}-\mathbf{u}_h]=0\quad\forall\,(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h, \tag{4.1a}$$

$$[\mathcal{B}_h((\mathbf{t},\boldsymbol{\sigma})-(\mathbf{t}_h,\boldsymbol{\sigma}_h)),\mathbf{v}_h]-[\mathcal{S}_h(\mathbf{u}-\mathbf{u}_h),\mathbf{v}_h]-[\mathcal{C}_h(\mathbf{u}-\mathbf{u}_h),\mathbf{v}_h]=0\quad\forall\,\mathbf{v}_h\in V_h. \tag{4.1b}$$

The following result establishes the Céa estimate for (2.5) and (2.11).

Lemma 4.1 *Assume that*

$$0\;<\;\tau\;<\;\min\left\{\frac{1}{\theta},\;\frac{1}{2}\left(\frac{C_{\mathrm{SM}}}{(1+C_{\mathrm{SM}})\theta+\widehat{C}_1}\right),\;\frac{1}{\vartheta}\right\},$$

with $\theta>0$ defined in Lemma 3.12 and

$$\vartheta\;:=\;2\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(\frac{\widehat{C}_1+2\widehat{C}_2^2}{C_{\mathrm{inf}}}\right)\;>\;0.$$

Let $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u})\in\mathbb{L}^2(\Omega)\times\mathbb{H}(\mathbf{div};\Omega)\times\mathbf{L}^2(\Omega)$ and $((\mathbf{t}_h,\boldsymbol{\sigma}_h),\mathbf{u}_h)\in H_h\times V_h$ be the unique solutions of (2.5) and (2.11), respectively. Then, there hold the Céa error estimates

$$\begin{aligned}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h}\;&\leq\;2\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(1+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right)\inf_{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}\\&+\left\{\frac{\|\mathcal{B}_h\|}{C_{\mathrm{SM}}}+\left\{1+\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right\}\left(\frac{C_{\mathrm{inf}}}{C_{\mathrm{LC}}}\right)\right\}\inf_{\mathbf{v}_h\in V_h}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega},\end{aligned} \tag{4.2}$$

and

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}\;&\leq\;2\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(1+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right)\inf_{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}\\&+2\left\{1+\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right\}\inf_{\mathbf{v}_h\in V_h}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}.\end{aligned} \tag{4.3}$$

Proof We proceed as in [33, Proposition 4.1]. In fact, we first set $H_h=\widetilde{H}_h\oplus\widetilde{H}_h^{\perp}$, with $\widetilde{H}_h$ being the kernel of $\mathcal{B}_h$. Hence, given $(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h$, we let $(\mathbf{r}_h,\boldsymbol{\rho}_h)\in\widetilde{H}_h^{\perp}$ be the unique solution of

$$[\mathcal{B}_h(\mathbf{r}_h, \boldsymbol{\rho}_h), \mathbf{v}_h] \;=\; [\mathcal{B}_h((\mathbf{t}, \boldsymbol{\sigma}) - (\mathbf{s}_h, \boldsymbol{\tau}_h)) - \mathcal{S}_h(\mathbf{u} - \mathbf{u}_h) - \mathcal{C}_h(\mathbf{u} - \mathbf{u}_h), \mathbf{v}_h] \quad \forall\, \mathbf{v}_h \in V_h,$$

which there exists thanks to the discrete inf-sup condition and the continuity of $\mathcal{B}_h$. Then, there holds

$$\begin{aligned} C_{\inf}\, \|(\mathbf{r}_h, \boldsymbol{\rho}_h)\|_{H_h} &\le \sup_{\substack{\mathbf{v}_h \in V_h \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{[\mathcal{B}_h(\mathbf{r}_h, \boldsymbol{\rho}_h), \mathbf{v}_h]}{\|\mathbf{v}_h\|_{0,\Omega}} \\ &= \sup_{\substack{\mathbf{v}_h \in V_h \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{[\mathcal{B}_h((\mathbf{t}, \boldsymbol{\sigma}) - (\mathbf{s}_h, \boldsymbol{\tau}_h)) - \mathcal{S}_h(\mathbf{u} - \mathbf{u}_h) - \mathcal{C}_h(\mathbf{u} - \mathbf{u}_h), \mathbf{v}_h]}{\|\mathbf{v}_h\|_{0,\Omega}} \\ &\le \|\mathcal{B}_h\|\, \|(\mathbf{t}, \boldsymbol{\sigma}) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h} + \{\, \|\mathcal{S}_h\| + \|\mathcal{C}_h\| \,\}\, \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \end{aligned}$$

that is

$$\|(\mathbf{r}_h, \boldsymbol{\rho}_h)\|_{H_h} \le \frac{\|\mathcal{B}_h\|}{C_{\inf}} \|(\mathbf{t}, \boldsymbol{\sigma}) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_{H_h} + \left\{ \frac{\|\mathcal{S}_h\| + \|\mathcal{C}_h\|}{C_{\inf}} \right\} \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}. \quad (4.4)$$

Also, note by construction of $(\mathbf{r}_h, \boldsymbol{\rho}_h) \in \widetilde{H}_h^{\perp}$ and (4.1b) that there holds

$$[\mathcal{B}_h((\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)), \mathbf{v}_h] = 0 \quad \forall\, \mathbf{v}_h \in V_h. \quad (4.5)$$

Next, applying the strong monotonicity of $\mathcal{A}_h$ and (4.1a), we get

$$\begin{aligned} &C_{\mathrm{SM}}\, \|(\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)\|_{H_h}^2 \\ &\quad \le [\mathcal{A}_h((\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h)) - \mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)] \\ &\quad = [\mathcal{A}_h((\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h)) - \mathcal{A}_h(\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)] \\ &\qquad + [\mathcal{A}_h(\mathbf{t}, \boldsymbol{\sigma}) - \mathcal{A}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)] \\ &\quad = [\mathcal{A}_h((\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h)) - \mathcal{A}_h(\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)] \\ &\qquad - [\mathcal{B}_h((\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)), \mathbf{u} - \mathbf{u}_h]. \end{aligned}$$

In turn, it follows from (4.5) that we can replace $\mathbf{u}_h$ by $\mathbf{v}_h \in V_h$ in the foregoing expression involving $\mathcal{B}_h$, and hence we obtain

$$\begin{aligned} &C_{\mathrm{SM}}\, \|(\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)\|_{H_h}^2 \\ &\quad \le [\mathcal{A}_h((\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h)) - \mathcal{A}_h(\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)] \\ &\qquad - [\mathcal{B}_h((\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)), \mathbf{u} - \mathbf{v}_h] \\ &\quad \le C_{\mathrm{LC}}\, \|(\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}, \boldsymbol{\sigma})\|_{H_h}\, \|(\mathbf{s}_h + \mathbf{r}_h, \boldsymbol{\tau}_h + \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)\|_{H_h} \\ &\qquad + \|\mathcal{B}_h\|\, \|(\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)\|_{H_h}\, \|\mathbf{u} - \mathbf{v}_h\|_{0,\Omega}, \end{aligned}$$

which yields

$$\begin{aligned} \|(\mathbf{s}_h, \boldsymbol{\tau}_h) + (\mathbf{r}_h, \boldsymbol{\rho}_h) - (\mathbf{t}_h, \boldsymbol{\sigma}_h)\|_{H_h} &\le \frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}} \|(\mathbf{t}, \boldsymbol{\sigma}) - (\mathbf{s}_h, \boldsymbol{\tau}_h) - (\mathbf{r}_h, \boldsymbol{\rho}_h)\|_{H_h} \\ &\quad + \frac{\|\mathcal{B}_h\|}{C_{\mathrm{SM}}} \|\mathbf{u} - \mathbf{v}_h\|_{0,\Omega}. \end{aligned}$$

Thus, by triangle inequality we deduce that

$$\begin{aligned}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h} &\leq \|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)-(\mathbf{r}_h,\boldsymbol{\rho}_h)\|_{H_h}\\ &\quad+\|(\mathbf{s}_h,\boldsymbol{\tau}_h)+(\mathbf{r}_h,\boldsymbol{\rho}_h)-(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h}\\ &\leq \left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)-(\mathbf{r}_h,\boldsymbol{\rho}_h)\|_{H_h}+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{SM}}}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}\\ &\leq \left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}+\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\|(\mathbf{r}_h,\boldsymbol{\rho}_h)\|_{H_h}\\ &\quad+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{SM}}}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega},\end{aligned}$$

which, together with (4.4) and the fact that $(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h$ and $\mathbf{v}_h\in V_h$ are arbitrary, imply

$$\begin{aligned}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h} &\leq \left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(1+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right)\inf_{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}\\ &\quad+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{SM}}}\inf_{\mathbf{v}_h\in V_h}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}+\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{\|\mathcal{S}_h\|+\|\mathcal{C}_h\|}{C_{\mathrm{inf}}}\right)\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}.\end{aligned}\tag{4.6}$$

On the other hand, using the inf-sup condition for $\mathcal{B}_h$, (4.1a), and the Lipschitz-continuity of $\mathcal{A}_h$, we find that for each $\mathbf{v}_h\in V_h$ there holds

$$\begin{aligned}C_{\mathrm{inf}}\,\|\mathbf{v}_h-\mathbf{u}_h\|_{0,\Omega} &\leq \sup_{\substack{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h\\(\mathbf{s}_h,\boldsymbol{\tau}_h)\neq\mathbf{0}}}\frac{[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{v}_h-\mathbf{u}_h]}{\|(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}}\\ &= \sup_{\substack{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h\\(\mathbf{s}_h,\boldsymbol{\tau}_h)\neq\mathbf{0}}}\frac{[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{v}_h-\mathbf{u}]+[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{u}-\mathbf{u}_h]}{\|(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}}\\ &= \sup_{\substack{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h\\(\mathbf{s}_h,\boldsymbol{\tau}_h)\neq\mathbf{0}}}\frac{[\mathcal{B}_h(\mathbf{s}_h,\boldsymbol{\tau}_h),\mathbf{v}_h-\mathbf{u}]-[\mathcal{A}_h(\mathbf{t},\boldsymbol{\sigma})-\mathcal{A}_h(\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{s}_h,\boldsymbol{\tau}_h)]}{\|(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}}\\ &\leq \|\mathcal{B}_h\|\,\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}+C_{\mathrm{LC}}\,\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h},\end{aligned}$$

which, together with an application of the triangle inequality, gives

$$\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}\leq\left(1+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right)\inf_{\mathbf{v}_h\in V_h}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}+\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\,\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{t}_h,\boldsymbol{\sigma}_h)\|_{H_h}.\tag{4.7}$$

Next, by substituting (4.6) into (4.7), we arrive at

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega} &\leq \left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(1+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right)\inf_{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}\\ &\quad+\left\{1+\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right\}\inf_{\mathbf{v}_h\in V_h}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}\\ &\quad+\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(\frac{\|\mathcal{S}_h\|+\|\mathcal{C}_h\|}{C_{\mathrm{inf}}}\right)\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}.\end{aligned}$$

In turn, we know from Lemma 3.11 that $\|\mathcal{S}_h\|\leq\tau\widehat{C}_1$ and $\|\mathcal{C}_h\|\leq 2\tau\widehat{C}_1^2$, and hence, recalling that $\tau<\frac{1}{\vartheta}$, we deduce that

$$\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(\frac{\|\mathcal{S}_h\|+\|\mathcal{C}_h\|}{C_{\mathrm{inf}}}\right)\leq\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(\frac{\widehat{C}_1+2\widehat{C}_1^2}{C_{\mathrm{inf}}}\right)\tau\;<\;\frac{1}{2},$$

which allows to conclude from the previous inequality that

$$\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega} \leq 2\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\left(\frac{C_{\mathrm{LC}}}{C_{\mathrm{inf}}}\right)\left(1+\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right)\inf_{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in H_h}\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_{H_h}$$
$$+2\left\{1+\left(1+\frac{C_{\mathrm{LC}}}{C_{\mathrm{SM}}}\right)\frac{\|\mathcal{B}_h\|}{C_{\mathrm{inf}}}\right\}\inf_{\mathbf{v}_h\in V_h}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}. \tag{4.8}$$

Finally, it is easy to see that (4.6) and (4.8) provide (4.2) and (4.3), thus finishing the proof. □

Next, in order to provide the rate of convergence of the discontinuous Galerkin scheme (2.11), we need the approximation properties of the finite element subspaces involved. For this purpose, given $T \in \mathcal{T}_h$, we let $\boldsymbol{\mathcal{P}}_T^k : \mathbb{L}^2(T) \to \mathbb{P}_k(T)$ and $\mathcal{P}_T^{k-1} : \mathbf{L}^2(T) \to \mathbf{P}_{k-1}(T)$ be the $\mathbb{L}^2(T)$ and $\mathbf{L}^2(T)-$ orthogonal projectors, respectively. It is well known (see, e.g. [7,18]) that for each $\mathbf{s} \in \mathbb{H}^\ell(T)$ and $\mathbf{v} \in \mathbf{H}^{\ell+1}(T)$ there holds

$$\|\mathbf{s}-\boldsymbol{\mathcal{P}}_T^k(\mathbf{s})\|_{0,T} \leq C h_T^{\min\{\ell,k+1\}}|\mathbf{s}|_{\ell,T} \quad \forall\, T \in \mathcal{T}_h, \tag{4.9}$$

and

$$\|\mathbf{v}-\mathcal{P}_T^{k-1}(\mathbf{v})\|_{0,T} \leq C h_T^{\min\{\ell+1,k\}}|\mathbf{v}|_{\ell+1,T} \quad \forall\, T \in \mathcal{T}_h. \tag{4.10}$$

On the other hand, let $\Pi_T^{k-1} : \mathbb{H}^1(T) \to \mathbb{P}_k(T)$ be the Raviart–Thomas interpolation operator (see [2,18,31]), which satisfies the approximation property

$$\|\boldsymbol{\tau}-\Pi_T^{k-1}(\boldsymbol{\tau})\|_{\mathbf{div},T} \leq C h_T^{\min\{\ell,k\}}\left\{|\boldsymbol{\tau}|_{\ell,T} + |\mathbf{div}(\boldsymbol{\tau})|_{\ell,T}\right\} \quad \forall\, T \in \mathcal{T}_h, \tag{4.11}$$

and for each $\boldsymbol{\tau} \in \mathbb{H}^\ell(T)$ such that $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{H}^\ell(T)$, with $\ell \geq 1$. Moreover, the interpolation operator Π_T^{k-1} can also be defined as a bounded linear operator from the larger space $\mathbb{H}^\ell(T) \cap \mathbb{H}(\mathbf{div};T)$ into $\mathbb{P}_k(T)$ for all $\ell \in (0,1]$ (see, e.g. [23, Theorem 3.16]). In this case there holds the following interpolation error estimate (see [18, Lemma 3.19])

$$\|\boldsymbol{\tau}-\Pi_T^{k-1}(\boldsymbol{\tau})\|_{0,T} \leq C h_T^{\ell}\left\{|\boldsymbol{\tau}|_{\ell,T} + \|\mathbf{div}(\boldsymbol{\tau})\|_{0,T}\right\} \quad \forall\, T \in \mathcal{T}_h,$$

which, together with (4.11), implies for $\ell > 0$ that

$$\|\boldsymbol{\tau}-\Pi_T^{k-1}(\boldsymbol{\tau})\|_{\mathbf{div},T} \leq C h_T^{\min\{\ell,k\}}\left\{|\boldsymbol{\tau}|_{\ell,T} + \|\mathbf{div}(\boldsymbol{\tau})\|_{\ell,T}\right\} \quad \forall\, T \in \mathcal{T}_h.$$

On the other hand, observe that, given $Z := \{\boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \boldsymbol{\tau}|_T \in \mathbb{H}^\ell(T) \quad \forall\, T \in \mathcal{T}_h\}$, we can define $\Pi_{\Sigma_h} : \mathbb{H}(\mathbf{div};\Omega) \cap Z \to \Sigma_h$ by

$$\Pi_{\Sigma_h}(\boldsymbol{\tau})|_T := \Pi_T^{k-1}(\boldsymbol{\tau}|_T) + d\,\mathbb{I} \quad \forall\, T \in \mathcal{T}_h,$$

with $d := -\frac{1}{n|\Omega|}\sum_{T\in\mathcal{T}_h}\int_T \mathrm{tr}\left(\Pi_T^{k-1}(\boldsymbol{\tau}|_T)\right) \in R$. Then, it is easy to prove that

$$\|\boldsymbol{\tau}-\Pi_{\Sigma_h}(\boldsymbol{\tau})\|_{\Sigma_h}^2 \leq \sum_{T\in\mathcal{T}_h}\|\boldsymbol{\tau}-\Pi_T^{k-1}(\boldsymbol{\tau})\|_{\mathbf{div},T}^2 \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)\cap Z,$$

and hence

$$\|\boldsymbol{\tau}-\Pi_{\Sigma_h}(\boldsymbol{\tau})\|_{\Sigma_h} \leq C\sum_{T\in\mathcal{T}_h} h_T^{\min\{\ell,k\}}\left\{|\boldsymbol{\tau}|_{\ell,T} + \|\mathbf{div}(\boldsymbol{\tau})\|_{\ell,T}\right\}. \tag{4.12}$$

In this way, as a consequence of (4.9), (4.10), (4.12), and the usual interpolation estimates, we find that S_h, Σ_h and V_h satisfy the following approximation properties:
($\mathbf{AP}_h^{\mathbf{t}}$) For each $\ell \geq 0$ and for each $\mathbf{s} \in \mathbb{H}^\ell(\Omega)$ there exists $\mathbf{s}_h \in S_h$ such that

$$\|\mathbf{s} - \mathbf{s}_h\|_{0,\Omega} \leq C \sum_{T \in \mathcal{T}_h} h_T^{\min\{\ell, k+1\}} |\mathbf{s}|_{\ell,T}.$$

($\mathbf{AP}_h^{\boldsymbol{\sigma}}$) For each $\ell > 0$ and for each $\boldsymbol{\tau} \in \mathbb{H}^\ell(\Omega)$ with $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{H}^\ell(\Omega)$ there exists $\boldsymbol{\tau}_h \in \Sigma_h$ such that

$$\|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\Sigma_h} \leq C \sum_{T \in \mathcal{T}_h} h_T^{\min\{\ell, k\}} \left\{ |\boldsymbol{\tau}|_{\ell,T} + \|\mathbf{div}(\boldsymbol{\tau})\|_{\ell,T} \right\}.$$

($\mathbf{AP}_h^{\mathbf{u}}$) For each $\ell \geq 0$ and for each $\mathbf{v} \in \mathbf{H}^\ell(\Omega)$ there exists $\mathbf{v}_h \in V_h$ such that

$$\|\mathbf{v} - \mathbf{v}_h\|_{0,\Omega} \leq C \sum_{T \in \mathcal{T}_h} h_T^{\min\{\ell+1, k\}} |\mathbf{v}|_{\ell+1,T}.$$

The following theorem establishes the theoretical rates of convergence of the discrete scheme (2.11), under suitable regularity assumptions on the exact solution.

Theorem 4.1 *Assume the same hypotheses of Lemma 4.1. In addition, suppose that there exists an integer $\ell > 0$ such that $\mathbf{t}|_T \in \mathbb{H}^\ell(T)$, $\boldsymbol{\sigma}|_T \in \mathbb{H}^\ell(T)$, $\mathbf{div}(\boldsymbol{\sigma}|_T) \in \mathbf{H}^\ell(T)$ and $\mathbf{u}|_T \in \mathbf{H}^{\ell+1}(T)$, for all $T \in \mathcal{T}_h$. Then, there exists $C > 0$, independent of h and the polynomial approximation degree k, such that*

$$\begin{aligned} &\|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Sigma_h} + \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \\ &\leq C \sum_{T \in \mathcal{T}_h} h_T^{\min\{\ell, k\}} \left\{ |\mathbf{t}|_{\ell,T} + |\boldsymbol{\sigma}|_{\ell,T} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{\ell,T} + |\mathbf{u}|_{\ell+1,T} \right\}. \end{aligned}$$

Proof It follows from the Céa estimate (cf. Lemma 4.1) and the approximation properties ($\mathbf{AP}_h^{\mathbf{t}}$), ($\mathbf{AP}_h^{\boldsymbol{\sigma}}$) and ($\mathbf{AP}_h^{\mathbf{u}}$). □

Note from the previous theorem and (3.7) that we can also conclude that

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} \leq C \sum_{T \in \mathcal{T}_h} h_T^{\min\{\ell, k\}} \left\{ |\mathbf{t}|_{\ell,T} + |\boldsymbol{\sigma}|_{\ell,T} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{\ell,T} + |\mathbf{u}|_{\ell+1,T} \right\}. \quad (4.13)$$

Furthermore, we know from (2.1) that $p = -\frac{1}{n}\mathrm{tr}\,(\boldsymbol{\sigma})$, which suggests to define the following postprocessed approximation of the pressure:

$$p_h := -\frac{1}{n}\,\mathrm{tr}\,(\boldsymbol{\sigma}_h) \quad \text{in} \quad \Omega,$$

and therefore

$$\|p - p_h\|_{0,\Omega} = \frac{1}{n}\|\mathrm{tr}\,(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{0,\Omega} \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega}, \quad (4.14)$$

which, thanks to (4.13), gives the a priori error estimate for the pressure.

Now, as in [11], we measure the errors of quantities defined on $\partial\mathcal{T}_h$ with the seminorm:

$$\|\boldsymbol{\mu}\|_h := \left\{ \sum_{T \in \mathcal{T}_h} h_T \|\boldsymbol{\mu}\|_{0,\partial T}^2 \right\}^{1/2},$$

and we let $\Pi_{\mathcal{E}_h} : \mathbf{L}^2(\mathcal{E}_h) \to \mathbf{P}_k(\mathcal{E}_h)$ be the orthogonal projection onto the space of piecewise polynomials of degree less than or equals to k on $\mathcal{E}_h$. Next, we end this section with the a priori error estimate for the trace of the velocity unknown, which is established next.

Theorem 4.2 *Assume the same hypotheses of Theorem 4.1. Then, there exists $C > 0$, independent of h and the polynomial approximation degree k, such that*

$$\|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \widehat{\mathbf{u}}_h\|_h \leq C \sum_{T\in\mathcal{T}_h} h_T^{\min\{\ell,k\}} \left\{ |\mathbf{t}|_{\ell,T} + |\boldsymbol{\sigma}|_{\ell,T} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{\ell,T} + |\mathbf{u}|_{\ell+1,T} \right\}.$$

Proof Since $\Pi_{\mathcal{E}_h}(\mathbf{u}) = \Pi_\Gamma(\mathbf{g}) = \widehat{\mathbf{u}}_h$ on $\mathcal{E}_h^\partial$, we only need to compute the error for each $F \in \mathcal{E}_h^i$. In fact, we have

$$\begin{aligned}
\|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \widehat{\mathbf{u}}_h\|_h^2 &= \sum_{T\in\mathcal{T}_h} \sum_{F\in\partial T\setminus\Gamma} h_T \|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \boldsymbol{\lambda}_h\|_{0,F}^2 \\
&\leq \widetilde{C} \sum_{T\in\mathcal{T}_h} \sum_{F\in\partial T\setminus\Gamma} \mathtt{h} \|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \boldsymbol{\lambda}_h\|_{0,F}^2 = 2\widetilde{C} \sum_{F\in\mathcal{E}_h^i} \mathtt{h} \|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \boldsymbol{\lambda}_h\|_{0,F}^2,
\end{aligned}$$

with $\widetilde{C} \geq 1$ depending only on the shape regularity of the mesh. Then, according to (2.10), (3.6) and the fact that $[\![\boldsymbol{\sigma}]\!] = \mathbf{0}$ on $\mathcal{E}_h^i$, we obtain that

$$\begin{aligned}
\|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \widehat{\mathbf{u}}_h\|_h^2 &\leq 2\widetilde{C} \sum_{F\in\mathcal{E}_h^i} \mathtt{h} \left\| \Pi_{\mathcal{E}_h}(\mathbf{u}) - \{\!\{\mathbf{u}_h\}\!\} + \frac{1}{2}(\tau\mathtt{h})^{-1} [\![\boldsymbol{\sigma}_h]\!] \right\|_{0,F}^2 \\
&\leq C \sum_{F\in\mathcal{E}_h^i} \left\{ \|\mathtt{h}^{1/2} \left(\Pi_{\mathcal{E}_h}(\mathbf{u}) - \{\!\{\mathbf{u}_h\}\!\} \right) \|_{0,F}^2 + \frac{1}{4\tau} \|(\tau\mathtt{h})^{-1/2} [\![\boldsymbol{\sigma} - \boldsymbol{\sigma}_h]\!] \|_{0,F}^2 \right\} \\
&\leq C \left\{ \|\mathtt{h}^{1/2} \left(\Pi_{\mathcal{E}_h}(\mathbf{u}) - \{\!\{\mathbf{u}_h\}\!\} \right) \|_{0,\mathcal{E}_h^i}^2 + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Sigma_h}^2 \right\} \\
&\leq C \left\{ \|\mathtt{h}^{1/2} \Pi_{\mathcal{E}_h}(\mathbf{u} - \boldsymbol{\mathcal{P}}_{1,h}^k(\mathbf{u}))\|_{0,\mathcal{E}_h^i}^2 + \|\mathtt{h}^{1/2} \{\!\{\boldsymbol{\mathcal{P}}_{1,h}^k(\mathbf{u}) - \mathbf{u}_h\}\!\}\|_{0,\mathcal{E}_h^i}^2 + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Sigma_h}^2 \right\},
\end{aligned} \tag{4.15}$$

where, denoting $\mathbf{X}_h^k := \left\{ \mathbf{v}_h \in \mathbf{C}(\bar{\Omega}) \ : \ \mathbf{v}_h|_T \in \mathbf{P}_k(T) \quad \forall\, T \in \mathcal{T}_h \right\}$, we let $\boldsymbol{\mathcal{P}}_{1,h}^k : \mathbf{H}^1(\Omega) \to \mathbf{X}_h^k$ be the orthogonal projector, which satisfies

$$\|\mathbf{v} - \boldsymbol{\mathcal{P}}_{1,h}^k(\mathbf{v})\|_{1,\Omega} \leq C_1 \sum_{T\in\mathcal{T}_h} h_T^{\min\{\ell,k\}} |\mathbf{v}|_{\ell+1,T} \quad \forall\, \mathbf{v} \in \mathbf{H}^{\ell+1}(T), \quad \forall\, T \in \mathcal{T}_h \tag{4.16}$$

and

$$\|\mathbf{v} - \boldsymbol{\mathcal{P}}_{1,h}^k(\mathbf{v})\|_{0,\Omega} \leq C_0 \sum_{T\in\mathcal{T}_h} h_T^{\min\{\ell+1,k+1\}} |\mathbf{v}|_{\ell+1,T} \quad \forall\, \mathbf{v} \in \mathbf{H}^{\ell+1}(T), \quad \forall\, T \in \mathcal{T}_h, \tag{4.17}$$

for $k \geq 1$ (see [18, Chapter 4] for details). Next, applying that $\|\Pi_{\mathcal{E}_h}\| \leq 1$ in the first term of (4.15), we find that

$$\begin{aligned}
\|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \widehat{\mathbf{u}}_h\|_h \leq C \Big\{ &\|\mathtt{h}^{1/2}(\mathbf{u} - \boldsymbol{\mathcal{P}}_{1,h}^k(\mathbf{u}))\|_{0,\mathcal{E}_h^i} \\
&+ \|\mathtt{h}^{1/2} \{\!\{\boldsymbol{\mathcal{P}}_{1,h}^k(\mathbf{u}) - \mathbf{u}_h\}\!\}\|_{0,\mathcal{E}_h^i} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Sigma_h} \Big\}.
\end{aligned}$$

Consequently, using (3.3) and the analogue of the part i) of Lemma 3.10 with $\mathbf{P}_k(\mathcal{T}_h)$ instead of V_h, we deduce that

$$\begin{aligned}\|\Pi_{\mathcal{E}_h}(\mathbf{u})-\widehat{\mathbf{u}}_h\|_h &\leq C\left\{\|\mathbf{u}-\mathcal{P}^k_{1,h}(\mathbf{u})\|_{1,\Omega}+\|\mathcal{P}^k_{1,h}(\mathbf{u})-\mathbf{u}_h\|_{0,\Omega}+\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\Sigma_h}\right\}\\ &\leq C\left\{\|\mathbf{u}-\mathcal{P}^k_{1,h}(\mathbf{u})\|_{1,\Omega}+\|\mathbf{u}-\mathcal{P}^k_{1,h}(\mathbf{u})\|_{0,\Omega}\right.\\ &\quad\left.+\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}+\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\Sigma_h}\right\},\end{aligned}$$

which, together with (4.16), (4.17) and Theorem 4.1, complete the proof. □

5 Implementation Considerations

In this section we describe some general aspects on the computational implementation of the discrete scheme proposed in Sect. 2. We remark that we refer to the original HDG system (2.9) since, as explained before, the equivalent reduced scheme given by (2.11) was introduced just for sake of the analysis. We begin by considering again problem (2.9) in a single element $T\in\mathcal{T}_h$ with Dirichlet's datum $\mathbf{g}=\mathbf{0}$ (as is usual, the boundary condition can be imposed later), that is

$$\begin{aligned}\int_T\boldsymbol{\psi}(\mathbf{t}_h):\mathbf{s}_h-\int_T\mathbf{s}_h:\boldsymbol{\sigma}_h^{\mathrm{d}}&=0,\\ \int_T\left\{\mathbf{t}_h-\kappa_1\boldsymbol{\psi}(\mathbf{t}_h)\right\}:\boldsymbol{\tau}_h^{\mathrm{d}}+\left\{\kappa_1\int_T\boldsymbol{\sigma}_h^{\mathrm{d}}:\boldsymbol{\tau}_h^{\mathrm{d}}+\kappa_2\int_T\mathbf{div}(\boldsymbol{\sigma}_h)\cdot\mathbf{div}(\boldsymbol{\tau}_h)\right\}&\\ +\int_T\mathbf{u}_h\cdot\mathbf{div}(\boldsymbol{\tau}_h)-\int_{\partial T}\boldsymbol{\tau}_h\boldsymbol{\nu}\cdot\boldsymbol{\lambda}_h&=-\kappa_2\int_T\mathbf{f}\cdot\mathbf{div}(\boldsymbol{\tau}_h),\\ -\int_T\mathbf{v}_h\cdot\mathbf{div}(\boldsymbol{\sigma}_h)+\int_{\partial T}\mathbf{S}\mathbf{u}_h\cdot\mathbf{v}_h-\int_{\partial T}\mathbf{S}\boldsymbol{\lambda}_h\cdot\mathbf{v}_h&=\int_T\mathbf{f}\cdot\mathbf{v}_h,\\ -\int_{\partial T}\boldsymbol{\sigma}_h\boldsymbol{\nu}\cdot\boldsymbol{\mu}_h+\int_{\partial T}\mathbf{S}\mathbf{u}_h\cdot\boldsymbol{\mu}_h-\int_{\partial T}\mathbf{S}\boldsymbol{\lambda}_h\cdot\boldsymbol{\mu}_h&=0,\end{aligned}$$

for all $(\mathbf{s}_h,\boldsymbol{\tau}_h,\mathbf{v}_h,\boldsymbol{\mu}_h)\in\mathbb{P}_k(T)\times\mathbb{P}_k(T)\times\mathbf{P}_{k-1}(T)\times\mathbf{P}_k(\partial T)$.

Note that, because of the null mean value condition of the trace of $\boldsymbol{\sigma}_h$, that is $\int_\Omega\mathrm{tr}(\boldsymbol{\sigma}_h)=0$, we can not establish the value of $\boldsymbol{\sigma}_h|_T$ only with the information from T (as it is natural in discontinuous Galerkin schemes). For that reason, and in order to rewrite the above local contribution in an equivalent form, we now define the local space

$$\Sigma_{h,0}(T):=\left\{\boldsymbol{\tau}\in\mathbb{P}_k(T):\int_T\mathrm{tr}(\boldsymbol{\tau})=0\right\},$$

for which there holds $\mathbb{P}_k(T)=\Sigma_{h,0}(T)\oplus P_0(T)\mathbb{I}$, where $\mathbb{I}\in R^{n\times n}$ is the identity matrix. Next, given $\boldsymbol{\sigma}_h,\boldsymbol{\tau}_h\in S_h$, we consider the local decomposition

$$\boldsymbol{\sigma}_h|_T=\widetilde{\boldsymbol{\sigma}}_h|_T+\rho_h|_T\,\mathbb{I}\quad\text{and}\quad\boldsymbol{\tau}_h|_T=\widetilde{\boldsymbol{\tau}}_h|_T+\zeta_h|_T\,\mathbb{I}\quad\forall\,T\in\mathcal{T}_h,$$

where $\widetilde{\boldsymbol{\sigma}}_h|_T,\ \widetilde{\boldsymbol{\tau}}_h|_T \in \Sigma_{h,0}(T),\ \rho_h|_T,\ \zeta_h|_T \in P_0(T)$, and rewrite the above local contribution as

$$
\begin{aligned}
\int_T \boldsymbol{\psi}(\mathbf{t}_h) : \mathbf{s}_h - \int_T \mathbf{s}_h : \widetilde{\boldsymbol{\sigma}}_h^{\mathrm{d}} &= 0,\\
\int_T \left\{\mathbf{t}_h - \kappa_1 \boldsymbol{\psi}(\mathbf{t}_h)\right\} : \widetilde{\boldsymbol{\tau}}_h^{\mathrm{d}} + \left\{\kappa_1 \int_T \widetilde{\boldsymbol{\sigma}}_h^{\mathrm{d}} : \widetilde{\boldsymbol{\tau}}_h^{\mathrm{d}} + \kappa_2 \int_T \mathbf{div}(\widetilde{\boldsymbol{\sigma}}_h)\cdot\mathbf{div}(\widetilde{\boldsymbol{\tau}}_h)\right\}&\\
+ \int_T \mathbf{u}_h \cdot \mathbf{div}(\widetilde{\boldsymbol{\tau}}_h) - \int_{\partial T} \widetilde{\boldsymbol{\tau}}_h \boldsymbol{\nu}\cdot\boldsymbol{\lambda}_h &= -\kappa_2 \int_T \mathbf{f}\cdot\mathbf{div}(\widetilde{\boldsymbol{\tau}}_h),\\
-\int_T \mathbf{v}_h \cdot \mathbf{div}(\widetilde{\boldsymbol{\sigma}}_h) + \int_{\partial T} \mathbf{S}\mathbf{u}_h\cdot\mathbf{v}_h - \int_{\partial T} \mathbf{S}\boldsymbol{\lambda}_h\cdot\mathbf{v}_h &= \int_T \mathbf{f}\cdot\mathbf{v}_h,\\
-\int_{\partial T} \widetilde{\boldsymbol{\sigma}}_h\boldsymbol{\nu}\cdot\boldsymbol{\mu}_h + \int_{\partial T} \mathbf{S}\mathbf{u}_h\cdot\boldsymbol{\mu}_h - \int_{\partial T} \mathbf{S}\boldsymbol{\lambda}_h\cdot\boldsymbol{\mu}_h - \int_{\partial T} \rho_h\boldsymbol{\mu}_h\cdot\boldsymbol{\nu} &= 0,\\
-\int_{\partial T} \zeta_h\boldsymbol{\lambda}_h\cdot\boldsymbol{\nu} &= 0\,,
\end{aligned}
$$

for all $(\mathbf{s}_h, \widetilde{\boldsymbol{\tau}}_h, \mathbf{v}_h, \boldsymbol{\mu}_h, \zeta_h) \in \mathbb{P}_k(T)\times\Sigma_{h,0}(T)\times\mathbf{P}_{k-1}(T)\times\mathbf{P}_k(\partial T)\times P_0(T)$. In addition, it is easy to see that the aforementioned condition on the trace of $\boldsymbol{\sigma}_h$ becomes

$$\sum_{T\in\mathcal{T}_h} \rho_h|_T\,|T| \;=\; 0,$$

which is imposed in the discrete system by means of a real Lagrange multiplier.

Then, applying the Newton–Raphson's method to the global nonlinear system, we translate the local contribution for the Newton's linear system in the mth iteration into the form

$$
\left(\begin{array}{ccccc:cc}
\mathbf{DA}_1(\mathbf{t}_h^m) & \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \delta\mathbf{t}_h^m & \mathbf{b}_1^m\\
-\mathbf{B}^T - \mathbf{DA}_2(\mathbf{t}_h^m) & \mathbf{H} & \mathbf{C} & -\mathbf{E} & \mathbf{0} & \delta\widetilde{\boldsymbol{\sigma}}_h^m & \mathbf{b}_2^m\\
\mathbf{0} & -\mathbf{C}^T & \mathbf{K} & -\mathbf{F} & \mathbf{0} & \delta\mathbf{u}_h^m & \mathbf{b}_3^m\\
\mathbf{0} & -\mathbf{E}^T & \mathbf{F}^T & -\mathbf{D} & \mathbf{G} & \delta\boldsymbol{\lambda}_h^m & \mathbf{b}_4^m\\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}^T & 0 & \delta\rho_h^m & \mathrm{b}_5^m
\end{array}\right),
$$

where $\delta\mathbf{t}_h^m$ corresponds to the mth update for the $\mathbf{t}_h$ variable, that is $\mathbf{t}_h^{m+1} = \mathbf{t}_h^m + \delta\mathbf{t}_h^m$, and similarly for the other variables. The discrete operators $\mathbf{DA}_i(\mathbf{r})$, $i\in\{1,2\}$, are the respective Gâteaux derivatives, given by

$$[\mathbf{DA}_1(\mathbf{r})\mathbf{t},\mathbf{s}] := \int_T \sum_{i,j,k,l=1}^{n} \frac{\partial}{\partial r_{kl}}\psi_{ij}(\mathbf{r})t_{kl}s_{ij} \;=\; \int_T \frac{\mu'(|\mathbf{r}|)}{|\mathbf{r}|}(\mathbf{r}:\mathbf{t})(\mathbf{r}:\mathbf{s}) \;+\; \int_T \mu(|\mathbf{r}|)\mathbf{t}:\mathbf{s},$$

and

$$[\mathbf{DA}_2(\mathbf{r})\mathbf{t},\mathbf{s}] := \kappa_1[\mathbf{DA}_1(\mathbf{r})\mathbf{t},\mathbf{s}^{\mathrm{d}}],$$

for all $\mathbf{r},\mathbf{t},\mathbf{s}\in\mathbb{L}^2(T)$, with $|\mathbf{r}| = \|\mathbf{r}\|_{R^{n\times n}} \neq 0$. In turn, using the same notation given in [6], the operators $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{H}$ are given as follows:

$$\mathbf{B} := \left[-\int_T \mathbf{s} : \boldsymbol{\tau}^{\mathrm{d}} \right] = -|\mathcal{J}_T| \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \otimes \mathbf{M},$$

$$\mathbf{C} := \left[\int_T \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\tau}) \right] = |\mathcal{J}_T| \, \mathbb{I} \otimes \left\{ \begin{pmatrix} \partial_x \hat{x} \\ \partial_y \hat{x} \end{pmatrix} \otimes \mathbf{DX} + \begin{pmatrix} \partial_x \hat{y} \\ \partial_y \hat{y} \end{pmatrix} \otimes \mathbf{DY} \right\},$$

and

$$\begin{aligned} \mathbf{H} &:= \left[\kappa_1 \int_T \boldsymbol{\sigma}^{\mathrm{d}} : \boldsymbol{\tau}^{\mathrm{d}} + \kappa_2 \int_T \mathbf{div}(\boldsymbol{\sigma}) \cdot \mathbf{div}(\boldsymbol{\tau}) \right] \\ &= -\kappa_1 \mathbf{B} + |\mathcal{J}_T| \kappa_2 \, \mathbb{I} \otimes \left\{ \begin{pmatrix} (\partial_x \hat{x})^2 & \partial_x \hat{x} \partial_y \hat{x} \\ \partial_x \hat{x} \partial_y \hat{x} & (\partial_y \hat{x})^2 \end{pmatrix} \otimes \mathbf{DXX} + \begin{pmatrix} \partial_x \hat{x} \partial_x \hat{y} & \partial_x \hat{x} \partial_y \hat{y} \\ \partial_y \hat{x} \partial_x \hat{y} & \partial_y \hat{x} \partial_y \hat{y} \end{pmatrix} \otimes \mathbf{DXY} \right. \\ &\quad \left. + \begin{pmatrix} \partial_x \hat{x} \partial_x \hat{y} & \partial_x \hat{x} \partial_y \hat{y} \\ \partial_y \hat{x} \partial_x \hat{y} & \partial_y \hat{x} \partial_y \hat{y} \end{pmatrix}^T \otimes \mathbf{DXY}^T + \begin{pmatrix} (\partial_x \hat{y})^2 & \partial_x \hat{y} \partial_y \hat{y} \\ \partial_x \hat{y} \partial_y \hat{y} & (\partial_y \hat{y})^2 \end{pmatrix} \otimes \mathbf{DYY} \right\}, \end{aligned}$$

where $\otimes$ is the Kronecker product, and given a basis $\{\hat{\varphi}_i\}$ of $P_k(\hat{T})$, $\mathbf{M} := \left[\int_{\hat{T}} \hat{\varphi}_i \hat{\varphi}_j\right]$ is the mass matrix, $\mathbf{DX} := \left[\int_{\hat{T}} \hat{\varphi}_j \partial_{\hat{x}} \hat{\varphi}_i\right]$, $\mathbf{DY} := \left[\int_{\hat{T}} \hat{\varphi}_j \partial_{\hat{y}} \hat{\varphi}_i\right]$, $\mathbf{DXX} := \left[\int_{\hat{T}} \partial_{\hat{x}} \hat{\varphi}_i \partial_{\hat{x}} \hat{\varphi}_j\right]$, $\mathbf{DXY} := \left[\int_{\hat{T}} \partial_{\hat{x}} \hat{\varphi}_i \partial_{\hat{y}} \hat{\varphi}_j\right]$, and $\mathbf{DYY} := \left[\int_{\hat{T}} \partial_{\hat{y}} \hat{\varphi}_i \partial_{\hat{y}} \hat{\varphi}_j\right]$, all them precomputed on the reference cell $\hat{T}$. In particular, when $\{\hat{\varphi}_i\}$ is the Dubiner basis [17], we only need to delete the first column in the above definition of $\mathbf{B}$, the first row in $\mathbf{C}$ and the first row and the first column in $\mathbf{H}$, in order to hold the belonging to the space $\Sigma_{h,0}(T)$. All the other discrete operators can be calculated similarly as in [6].

It is important to note here that the local submatrix

$$\begin{pmatrix} \mathbf{DA}_1(\mathbf{t}_h^m) & \mathbf{B} & \mathbf{0} \\ -\mathbf{B}^T - \mathbf{DA}_2(\mathbf{t}_h^m) & \mathbf{H} & \mathbf{C} \\ \mathbf{0} & -\mathbf{C}^T & \mathbf{K} \end{pmatrix} \in R^{(n^2 d_q + (n^2 d_q - 1) + n d_u) \times (n^2 d_q + (n^2 d_q - 1) + n d_u)},$$

with $d_q := \dim P_k(T)$ and $d_u := \dim P_{k-1}(T)$, is invertible when $\mu > 0$ and $|\mathbf{t}_h^m| \neq 0$. Then, as it is usual in the HDG methods, we can obtain the values of $\delta \mathbf{t}_h^m|_T$, $\delta \widetilde{\boldsymbol{\sigma}}_h^m|_T$ and $\delta \mathbf{u}_h^m|_T$ as functions of $\delta \boldsymbol{\lambda}_h^m|_T$ and $\delta \rho_h^m|_T$ (actually, they only depend on $\delta \boldsymbol{\lambda}_h^m|_T$). In other words, we can reduce the stencil of the global linear system on each iteration of the Newton's method.

Finally, we let

$$N_{\mathrm{total}} := (n^2 d_q + n^2 d_q + n d_u) \times (\#\text{ of element in } \mathcal{T}_h) + (n d_l) \times (\#\text{ of faces in } \mathcal{T}_h),$$

with $d_l := \dim P_k(F)$, $F \in \partial T$, be the total number of degrees of freedom (without including those for the pressure). In other words, N_{total} is the total number of unknowns defining $\mathbf{t}_h$, $\boldsymbol{\sigma}_h$, $\mathbf{u}_h$ and $\boldsymbol{\lambda}_h$. On the other hand, we let

$$N_{\mathrm{comp}} := (n d_l) \times (\#\text{ of faces in } \mathcal{T}_h) + (\#\text{ of element in } \mathcal{T}_h) + 1$$

be the number of degrees of freedom effectively employed in the computations, i.e, the total number of unknowns defining $\boldsymbol{\lambda}_h$, ρ_h and the Lagrange multiplier.

6 Numerical Results

In this section we present several numerical experiments illustrating the performance of the augmented HDG method introduced in Sect. 2. We set $\tau = 10^{-1}$ for each one of the

Table 1 History of convergence for Example 1

k	h	$\|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\Sigma_h}$		$\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}$		$\|\Pi_{\mathcal{E}_h}(\mathbf{u})-\widehat{\mathbf{u}}_h\|_h$		$\|p-p_h\|_{0,\Omega}$	
		Error	Order	Error	Order	Error	Order	Error	Order	Error	Order	Error	Order
1	0.2000	1.14e−0	–	3.03e−1	–	5.11e−0	–	4.77e−1	–	9.21e−2	–	1.98e−1	–
	0.1333	5.27e−1	1.91	1.40e−1	1.90	3.51e−0	0.93	3.19e−1	1.00	3.49e−2	2.39	9.18e−2	1.90
	0.1000	3.01e−1	1.95	8.01e−2	1.95	2.66e−0	0.96	2.39e−1	1.00	1.81e−2	2.28	5.25e−2	1.95
	0.0800	1.94e−1	1.97	5.16e−2	1.97	2.14e−0	0.98	1.91e−1	1.00	1.11e−2	2.20	3.38e−2	1.97
	0.0667	1.35e−1	1.98	3.60e−2	1.98	1.79e−0	0.98	1.59e−1	1.00	7.49e−3	2.15	2.36e−2	1.98
	0.0571	9.97e−2	1.98	2.65e−2	1.98	1.53e−0	0.99	1.37e−1	1.00	5.41e−3	2.11	1.74e−2	1.98
	0.0500	7.65e−2	1.99	2.03e−2	1.99	1.34e−0	0.99	1.20e−1	1.00	4.10e−3	2.09	1.33e−2	1.99
2	0.2000	9.17e−2	–	1.91e−2	–	5.91e−1	–	6.83e−2	–	6.40e−3	–	1.18e−2	–
	0.1333	2.87e−2	2.86	5.97e−3	2.86	2.74e−1	1.90	3.02e−2	2.01	1.39e−3	3.77	3.70e−3	2.87
	0.1000	1.24e−2	2.91	2.58e−3	2.92	1.56e−1	1.95	1.70e−2	2.00	4.58e−4	3.85	1.60e−3	2.93
	0.0800	6.45e−3	2.94	1.33e−3	2.95	1.01e−1	1.97	1.09e−2	2.00	1.92e−4	3.89	8.26e−4	2.95
	0.0667	3.77e−3	2.95	7.77e−4	2.96	7.02e−2	1.98	7.54e−3	2.00	9.44e−5	3.91	4.81e−4	2.97
	0.0571	2.39e−3	2.96	4.92e−4	2.97	5.17e−2	1.98	5.54e−3	2.00	5.15e−5	3.92	3.04e−4	2.97
	0.0500	1.61e−3	2.96	3.31e−4	2.98	3.97e−2	1.99	4.24e−3	2.00	3.05e−5	3.93	2.04e−4	2.98
3	0.2000	5.81e−3	–	1.20e−3	–	5.04e−2	–	7.31e−3	–	2.57e−4	–	7.45e−4	–
	0.1333	1.20e−3	3.89	2.51e−4	3.87	1.56e−2	2.89	2.17e−3	3.00	3.64e−5	4.82	1.56e−4	3.86
	0.1000	3.87e−4	3.94	8.10e−5	3.93	6.71e−3	2.94	9.15e−4	3.00	8.91e−6	4.89	5.03e−5	3.93
	0.0800	1.60e−4	3.96	3.35e−5	3.95	3.46e−3	2.96	4.69e−4	3.00	2.97e−6	4.92	2.08e−5	3.95
	0.0667	7.76e−5	3.97	1.63e−5	3.97	2.01e−3	2.97	2.71e−4	3.00	1.21e−6	4.94	1.01e−5	3.97
	0.0571	4.20e−5	3.98	8.81e−6	3.98	1.27e−3	2.98	1.71e−4	3.00	5.62e−7	4.95	5.48e−6	3.98
	0.0500	2.47e−5	3.98	5.18e−6	3.98	8.53e−4	2.99	1.14e−4	3.00	2.90e−7	4.96	3.22e−6	3.98

Table 2 History of convergence for Example 2

k	h	$\|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\Sigma_h}$		$\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}$		$\|\Pi_{\mathcal{E}_h}(\mathbf{u})-\widehat{\mathbf{u}}_h\|_h$		$\|p-p_h\|_{0,\Omega}$	
		Error	Order	Error	Order	Error	Order	Error	Order	Error	Order	Error	Order
1	0.2000	5.52e−1	–	5.85e−1	–	1.04e+1	–	4.75e−1	–	5.59e−2	–	3.37e−1	–
	0.1333	2.48e−1	1.97	2.64e−1	1.96	7.01e−0	0.98	3.17e−1	0.99	2.35e−2	2.14	1.53e−1	1.95
	0.1000	1.41e−1	1.98	1.50e−1	1.98	5.28e−0	0.99	2.38e−1	1.00	1.30e−2	2.07	8.64e−2	1.98
	0.0800	9.02e−2	1.99	9.60e−2	1.99	4.23e−0	0.99	1.91e−1	1.00	8.19e−3	2.06	5.55e−2	1.99
	0.0667	6.28e−2	1.99	6.68e−2	1.99	3.53e−0	0.99	1.59e−1	1.00	5.65e−3	2.04	3.86e−2	1.99
	0.0571	4.62e−2	1.99	4.91e−2	1.99	3.03e−0	1.00	1.36e−1	1.00	4.13e−3	2.03	2.84e−2	1.99
	0.0500	3.54e−2	1.99	3.76e−2	1.99	2.65e−0	1.00	1.19e−1	1.00	3.15e−3	2.02	2.17e−2	1.99
2	0.2000	4.77e−2	–	4.06e−2	–	1.31e−0	–	5.95e−2	–	2.62e−3	–	1.92e−2	–
	0.1333	1.44e−2	2.96	1.22e−2	2.96	5.90e−1	1.97	2.64e−2	2.00	5.68e−4	3.77	5.78e−3	2.96
	0.1000	6.10e−3	2.97	5.19e−3	2.98	3.30e−1	2.02	1.49e−2	2.00	1.86e−4	3.89	2.45e−3	2.98
	0.0800	3.15e−3	2.96	2.69e−3	2.95	2.14e−1	1.93	9.50e−3	2.00	7.99e−5	3.78	1.27e−3	2.96
	0.0667	1.83e−3	2.98	1.56e−3	2.98	1.49e−1	1.99	6.60e−3	2.00	3.93e−5	3.89	7.37e−4	2.98
	0.0571	1.16e−3	2.98	9.85e−4	2.98	1.10e−1	1.99	4.85e−3	2.00	2.15e−5	3.91	4.65e−4	2.99
	0.0500	7.76e−4	2.98	6.61e−4	2.99	8.40e−2	2.00	3.71e−3	2.00	1.27e−5	3.93	3.12e−4	2.99
3	0.2000	3.58e−3	–	3.24e−3	–	1.47e−1	–	5.35e−3	–	1.34e−4	–	1.36e−3	–
	0.1333	7.31e−4	3.92	6.95e−4	3.80	4.88e−2	2.72	1.58e−3	3.01	2.11e−5	4.56	2.95e−4	3.77
	0.1000	2.41e−4	3.87	2.30e−4	3.85	2.28e−2	2.64	6.64e−4	3.01	6.11e−6	4.30	9.64e−5	3.89
	0.0800	9.83e−5	4.01	9.41e−5	4.00	1.10e−2	3.27	3.39e−4	3.00	1.82e−6	5.42	3.97e−5	3.98
	0.0667	4.79e−5	3.95	4.58e−5	3.96	6.42e−3	2.96	1.96e−4	3.00	7.53e−7	4.85	1.92e−5	3.98
	0.0571	2.60e−5	3.96	2.48e−5	3.97	4.06e−3	2.97	1.24e−4	3.00	3.55e−7	4.87	1.04e−5	4.00
	0.0500	1.53e−5	3.95	1.46e−5	3.97	2.74e−3	2.95	8.28e−5	3.00	1.86e−7	4.85	6.07e−6	4.01

Table 3 Example 2, some errors for different values of τ

h	k	$\tau = 10^{-1}$	$\tau = 10^{0}$	$\tau = 10^{1}$	$\tau = 10^{2}$	$\tau = 10^{3}$
$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega}$						
0.0571	1	4.9118e−2	6.5855e−2	2.8347e−1	2.3974e−0	1.7163e+1
0.0571	2	9.8544e−4	1.3358e−3	6.2561e−3	5.5194e−2	3.8630e−1
0.0667	3	4.5751e−5	5.5785e−5	2.3936e−4	2.1504e−3	1.5109e−2
0.0667	4	2.9174e−6	2.9876e−6	5.7542e−6	4.6312e−5	3.4693e−4
$\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}$						
0.0571	1	1.3626e−1	1.3649e−1	1.3717e−1	1.5542e−1	2.7216e−1
0.0571	2	4.8468e−3	5.2903e−3	5.7489e−3	5.8752e−3	6.8390e−3
0.0667	3	1.9632e−4	2.3910e−4	3.0923e−4	3.3322e−4	3.7520e−4
0.0667	4	4.6136e−6	5.8874e−6	7.1539e−6	7.4307e−6	7.8061e−6
$\|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \widehat{\mathbf{u}}_h\|_h$						
0.0571	1	4.1323e−3	6.8604e−3	3.3457e−2	2.3303e−1	7.6095e−1
0.0571	2	2.1471e−5	3.8337e−5	2.0965e−4	1.8313e−3	1.2478e−2
0.0667	3	7.5312e−7	1.1598e−6	6.2942e−6	5.6190e−5	3.7855e−4
0.0667	4	3.8092e−8	4.0211e−8	1.0500e−7	9.0616e−7	6.6465e−6

4 examples to be reported, which, as shown below, works fine in all the cases. An a priori verification of the hypotheses on τ in Lemma 4.1 would certainly require the explicit knowledge of all the constants involved, which, however, is rarely possible. On the other hand, we take the stabilization parameter $\kappa_1 = \frac{\alpha_0}{\gamma_0^2}$, which obviously satisfies the assumption $\kappa_1 \in \left(0, \frac{2\alpha_0}{\gamma_0^2}\right)$ in Lemma 3.6, and then, as suggested by the value of the strong monotonicity constant C_{SM} at the end of its proof, we simply choose $\kappa_2 = \frac{\kappa_1}{2}$. The corresponding nonlinear algebraic system arising from (2.9) is solved by the Newton method with a tolerance of 10^{-6} and taking as initial iteration the solution of the associated linear Stokes problem (four iterations were required to achieve the given tolerance in each example). Now, according to the definitions given in Sect. 5, we recall that N_{total} is the total number of degrees of freedom, and N_{comp} is the number of degrees of freedom involved in the implementation of the Newton's method. To this respect, and even though we understand that a meaningful comparison makes sense between the N_{comp} of two different methods, in Example 4 below we display the information concerning N_{comp} and N_{total} only to appreciate the reduction in the degrees of freedom provided by our method, which is one of the key aspects of the HDG approaches. We do not perform any comparison with other method since we are not aware of another HDG type procedure dealing with our nonlinear problem.

The numerical results presented below were obtained using a C++ code, which was developed following the same techniques from [6]. In turn, the linear systems are solved using the conjugate gradient method with a relative tolerance of 10^{-6}.

In Example 1 we follow [11,30] and consider the linear Stokes problem given by the flow uncovered by Kovaszany [26]. This means that $\Omega := (-0.5, 1.5) \times (0, 2)$, $\mu = 0.1$, and the data $\mathbf{f}$ and $\mathbf{g}$ are chosen so that the exact solution is given by

Table 4 History of convergence for Example 3

k	h	$\|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\Sigma_h}$		$\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}$		$\|\Pi_{\mathcal{E}_h}(\mathbf{u})-\widehat{\mathbf{u}}_h\|_h$		$\|p-p_h\|_{0,\Omega}$	
		Error	Order	Error	Order	Error	Order	Error	Order	Error	Order	Error	Order
1	0.1667	8.54e−2	–	9.75e−2	–	7.65e−0	–	6.75e−2	–	1.04e−2	–	5.45e−2	–
	0.1111	6.59e−2	0.64	6.95e−2	0.84	8.62e−0	-0.30	4.51e−2	0.99	7.16e−3	0.92	3.72e−2	0.94
	0.0833	5.47e−2	0.64	5.46e−2	0.84	9.39e−0	-0.30	3.39e−2	1.00	5.43e−3	0.96	2.83e−2	0.96
	0.0667	4.74e−2	0.65	4.54e−2	0.83	1.00e+1	-0.30	2.71e−2	1.00	4.34e−3	1.00	2.29e−2	0.95
	0.0556	4.21e−2	0.65	3.91e−2	0.82	1.06e+1	-0.30	2.26e−2	1.00	3.60e−3	1.03	1.93e−2	0.95
	0.0455	3.69e−2	0.65	3.33e−2	0.81	1.13e+1	-0.30	1.85e−2	1.00	2.92e−3	1.05	1.60e−2	0.93
	0.0400	3.39e−2	0.65	3.01e−2	0.80	1.17e+1	-0.30	1.63e−2	1.00	2.54e−3	1.07	1.42e−2	0.92
2	0.1667	6.09e−2	–	5.32e−2	–	6.84e−0	–	2.45e−3	–	5.47e−3	–	2.23e−2	–
	0.1111	4.67e−2	0.65	3.91e−2	0.76	7.72e−0	-0.30	1.33e−3	1.50	3.09e−3	1.41	1.56e−2	0.87
	0.0833	3.87e−2	0.66	3.16e−2	0.74	8.41e−0	-0.30	8.65e−4	1.50	2.07e−3	1.39	1.23e−2	0.82
	0.0667	3.34e−2	0.66	2.69e−2	0.73	8.99e−0	-0.30	6.21e−4	1.48	1.52e−3	1.38	1.03e−2	0.80
	0.0556	2.96e−2	0.66	2.36e−2	0.72	9.50e−0	-0.30	4.75e−4	1.47	1.19e−3	1.37	8.96e−3	0.78
	0.0455	2.60e−2	0.66	2.04e−2	0.72	1.01e+1	-0.30	3.55e−4	1.46	9.03e−4	1.36	7.68e−3	0.77
	0.0400	2.39e−2	0.66	1.86e−2	0.72	1.05e+1	-0.30	2.95e−4	1.45	7.60e−4	1.35	6.97e−3	0.76
3	0.1667	4.48e−2	–	3.64e−2	–	5.99e−0	–	7.05e−4	–	2.36e−3	–	1.30e−2	–
	0.1111	3.43e−2	0.66	2.73e−2	0.72	6.76e−0	-0.30	3.86e−4	1.48	1.29e−3	1.49	9.57e−3	0.76
	0.0833	2.84e−2	0.66	2.22e−2	0.71	7.37e−0	-0.30	2.55e−4	1.45	8.44e−4	1.47	7.73e−3	0.74
	0.0667	2.45e−2	0.66	1.90e−2	0.71	7.89e−0	-0.30	1.85e−4	1.42	6.12e−4	1.45	6.57e−3	0.73
	0.0556	2.17e−2	0.66	1.67e−2	0.71	8.33e−0	-0.30	1.43e−4	1.41	4.71e−4	1.43	5.75e−3	0.73
	0.0455	1.91e−2	0.66	1.45e−2	0.70	8.85e−0	-0.30	1.09e−4	1.39	3.55e−4	1.41	4.98e−3	0.72
	0.0400	1.75e−2	0.66	1.32e−2	0.70	9.20e−0	-0.30	9.11e−5	1.37	2.97e−4	1.40	4.54e−3	0.72

Table 5 History of convergence for Example 4

k	h	N_{total}	N_{comp}	$\|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{0,\Omega}$		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\Sigma_h}$	
				Error	Order	Error	Order	Error	Order
1	0.3464	71100	15601	4.30e−1	–	4.35e−1	–	7.21e−0	–
	0.2474	194040	41749	2.22e−1	1.96	2.30e−1	1.88	5.19e−0	0.98
	0.1925	411156	87481	1.35e−1	1.98	1.43e−1	1.91	4.06e−0	0.98
	0.1732	563400	119401	1.10e−1	2.00	1.16e−1	1.97	3.65e−0	1.01
	0.1332	1235052	259585	6.52e−2	1.98	6.94e−2	1.95	2.82e−0	0.98
	0.1083	2299392	480769	4.32e−2	1.99	4.60e−2	1.97	2.29e−0	0.99
	0.0962	3271752	682345	3.41e−2	2.00	3.64e−2	1.99	2.04e−0	1.00
2	0.3464	173700	30451	4.43e−2	–	3.94e−2	–	1.15e−0	–
	0.2474	474516	81439	1.70e−2	2.85	1.49e−2	2.89	6.17e−1	1.86
	0.1925	1006020	170587	8.14e−3	2.92	7.13e−3	2.94	3.77e−1	1.96
	0.1732	1378800	232801	6.15e−3	2.66	5.51e−3	2.44	3.27e−1	1.34
	0.1332	3023748	505987	2.78e−3	3.03	2.44e−3	3.10	1.86e−1	2.15
	0.1083	5630976	936961	1.50e−3	2.97	1.31e−3	3.01	1.21e−1	2.07
	0.0962	8013168	1329697	1.06e−3	2.92	9.24e−4	2.95	9.61e−2	1.96
3	0.3464	342000	50251	5.96e−3	–	6.29e−3	–	2.55e−1	–
	0.2474	934920	134359	1.92e−3	3.38	2.13e−3	3.22	1.16e−1	2.33
	0.1925	1982880	281395	8.53e−4	3.22	9.92e−4	3.04	6.79e−2	2.14
	0.1732	2718000	384001	5.33e−4	4.47	6.75e−4	3.65	4.75e−2	3.38
	0.1332	5962320	834523	1.88e−4	3.97	2.41e−4	3.94	2.16e−2	3.00
	0.1083	11105280	1545217	8.22e−5	3.98	1.05e−4	3.99	1.16e−2	2.99
	0.0962	15804720	2192833	5.15e−5	3.96	6.61e−5	3.93	8.18e−3	2.98
k	h	N_{total}	N_{comp}	$\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}$		$\|\Pi_{\mathcal{E}_h}(\mathbf{u})-\widehat{\mathbf{u}}_h\|_h$		$\|p-p_h\|_{0,\Omega}$	
				Error	Order	Error	Order	Error	Order
1	0.3464	71100	15601	2.64e−1	–	1.66e−1	–	1.86e−1	–
	0.2474	194040	41749	1.89e−1	0.99	8.54e−2	1.98	1.01e−1	1.82
	0.1925	411156	87481	1.48e−1	0.99	5.18e−2	1.99	6.29e−2	1.87
	0.1732	563400	119401	1.33e−1	1.00	4.20e−2	2.00	5.12e−2	1.95
	0.1332	1235052	259585	1.02e−1	1.00	2.49e−2	1.99	3.08e−2	1.94
	0.1083	2299392	480769	8.31e−2	1.00	1.65e−2	2.00	2.05e−2	1.97
	0.0962	3271752	682345	7.39e−2	1.00	1.30e−2	2.00	1.62e−2	1.98
2	0.3464	173700	30451	4.03e−2	–	6.07e−3	–	1.33e−2	–
	0.2474	474516	81439	2.07e−2	1.98	1.75e−3	3.70	4.61e−3	3.14
	0.1925	1006020	170587	1.26e−2	1.99	6.66e−4	3.84	2.13e−3	3.08
	0.1732	1378800	232801	1.02e−2	1.99	4.72e−4	3.26	1.60e−3	2.72
	0.1332	3023748	505987	6.04e−3	1.99	1.62e−4	4.07	6.95e−4	3.18
	0.1083	5630976	936961	3.99e−3	2.00	7.09e−5	3.99	3.65e−4	3.10
	0.0962	8013168	1329697	3.15e−3	2.00	4.48e−5	3.89	2.55e−4	3.04
3	0.3464	342000	50251	4.39e−3	–	6.88e−4	–	1.63e−3	–
	0.2474	934920	134359	1.61e−3	2.98	1.62e−4	4.31	4.96e−4	3.54
	0.1925	1982880	281395	7.60e−4	2.99	5.67e−5	4.16	2.25e−4	3.14
	0.1732	2718000	384001	5.55e−4	2.99	3.20e−5	5.44	1.43e−4	4.31
	0.1332	5962320	834523	2.53e−4	2.99	8.67e−6	4.97	5.02e−5	3.99
	0.1083	11105280	1545217	1.36e−4	3.01	3.08e−6	4.98	2.20e−5	3.97
	0.0962	15804720	2192833	9.53e−5	3.00	1.72e−6	4.96	1.39e−5	3.93

$$\mathbf{u}(\mathbf{x}) = \left(1-\exp(\lambda x_1)\cos(2\pi x_2),\ \frac{\lambda}{2\pi}\exp(\lambda x_1)\sin(2\pi x_2)\right),$$
$$p(\mathbf{x}) = \frac{1}{2}\exp(2\lambda x_1) - \frac{1}{8\lambda}\left\{\exp(3\lambda)-\exp(-\lambda)\right\},$$

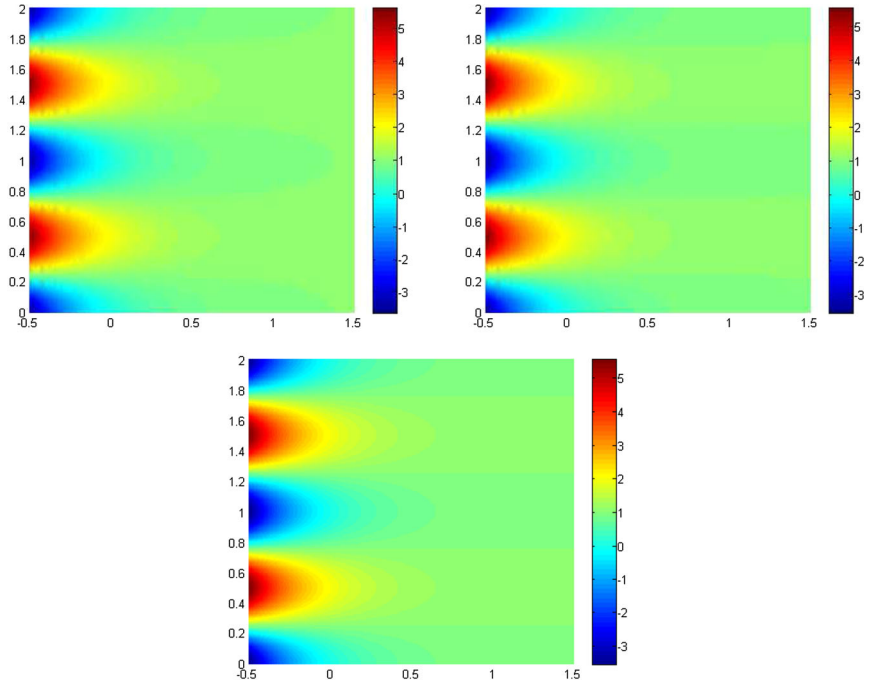

Fig. 1 Example 2, $u_{h,1}$ for $k = 2$ (*top-left*), for $k = 3$ (*top-right*), and its exact value (*bottom*)

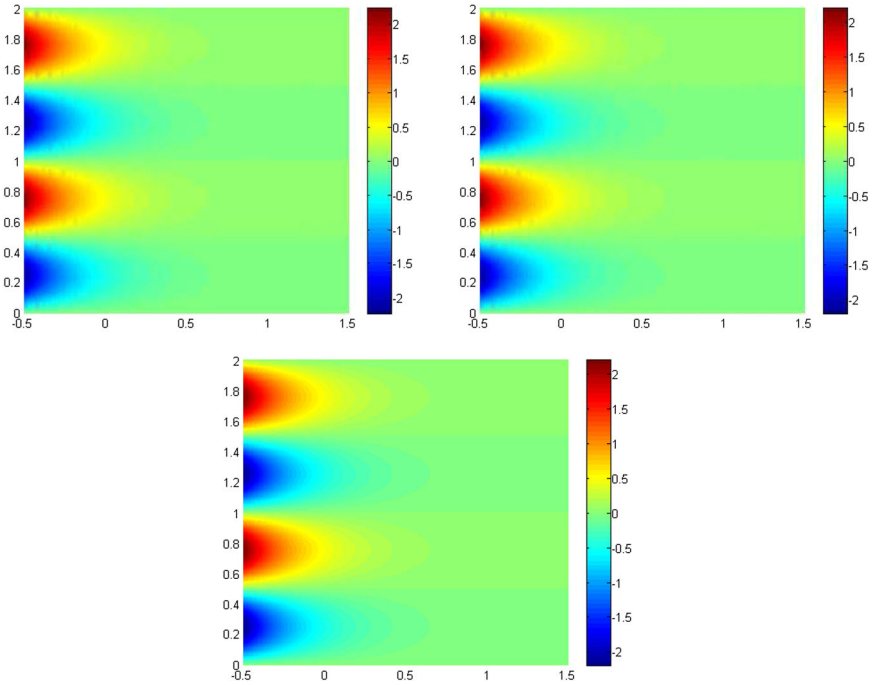

Fig. 2 Example 2, $u_{h,2}$ for $k = 2$ (*top-left*), for $k = 3$ (*top-right*), and its exact value (*bottom*)

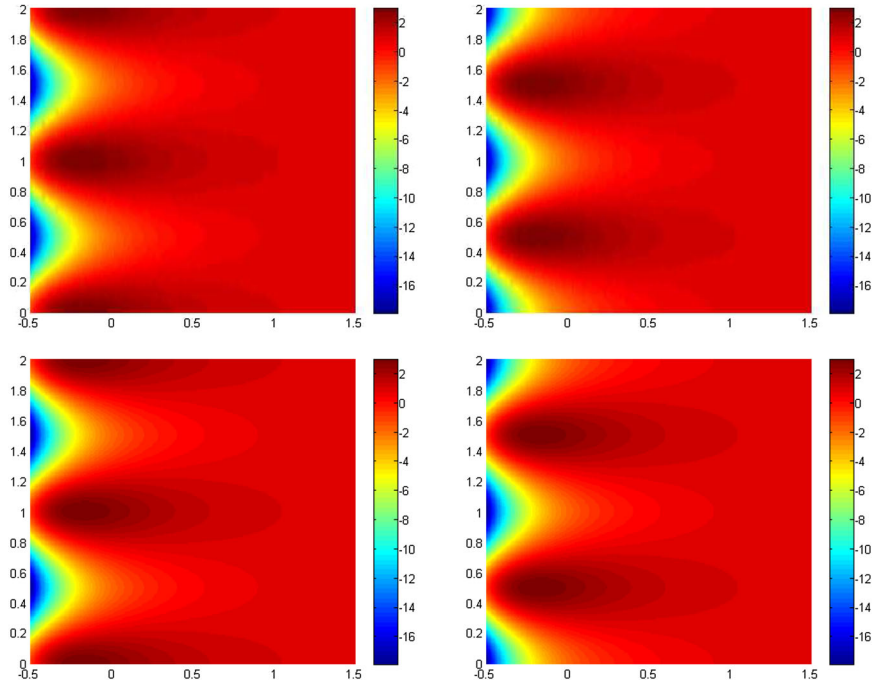

Fig. 3 Example 2, $\sigma_{h,11}$ (*top-left*) $\sigma_{h,22}$ (*top-right*) for $k = 2$, and its exact values (*bottom*)

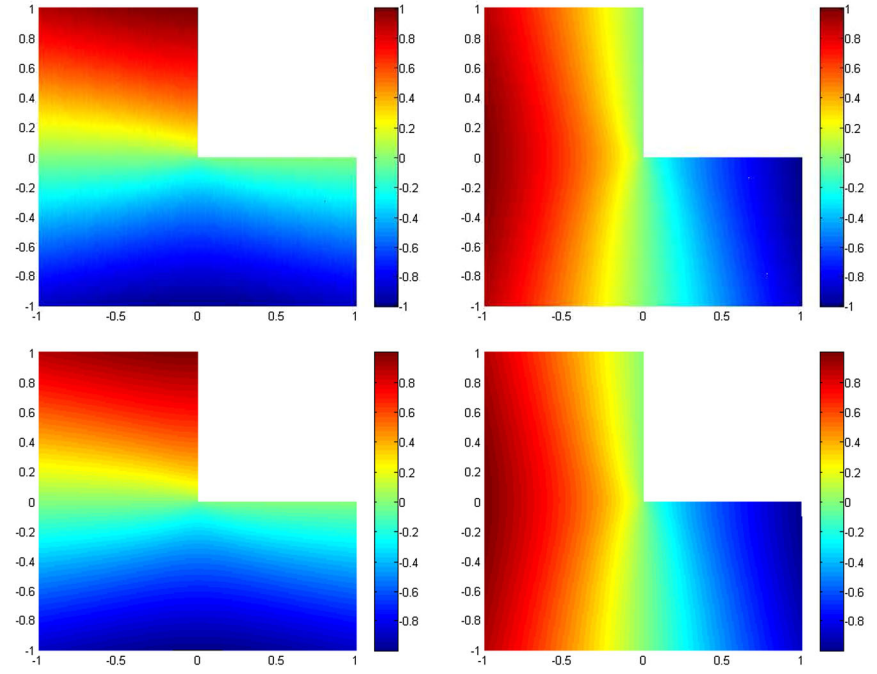

Fig. 4 Example 3, $u_{h,1}$ (*top-left*) and $u_{h,2}$ (*top-right*) for $k = 2$, and its exact values (*bottom*)

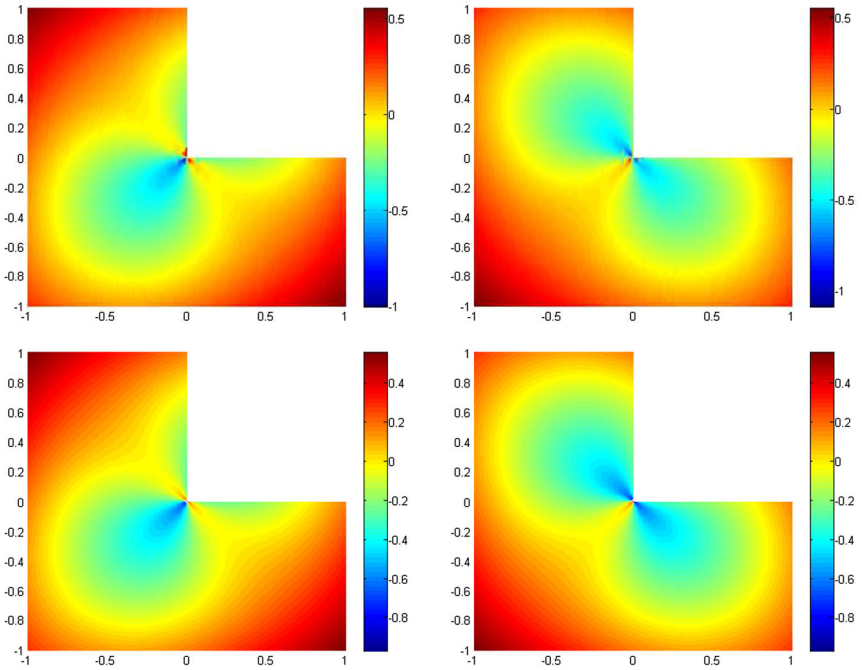

Fig. 5 Example 3, $\sigma_{h,11}$ (*top-left*) and $\sigma_{h,22}$ (*top-right*) for $k = 2$, and its exact values (*bottom*)

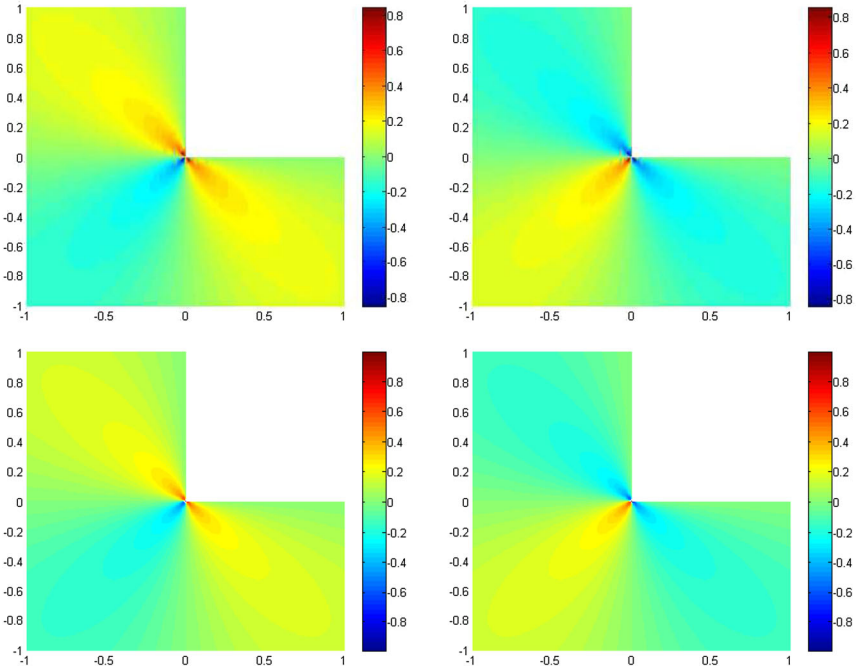

Fig. 6 Example 3, $t_{h,11}$ (*top-left*) and $t_{h,22}$ (*top-right*) for $k = 2$, and its exact values (*bottom*)

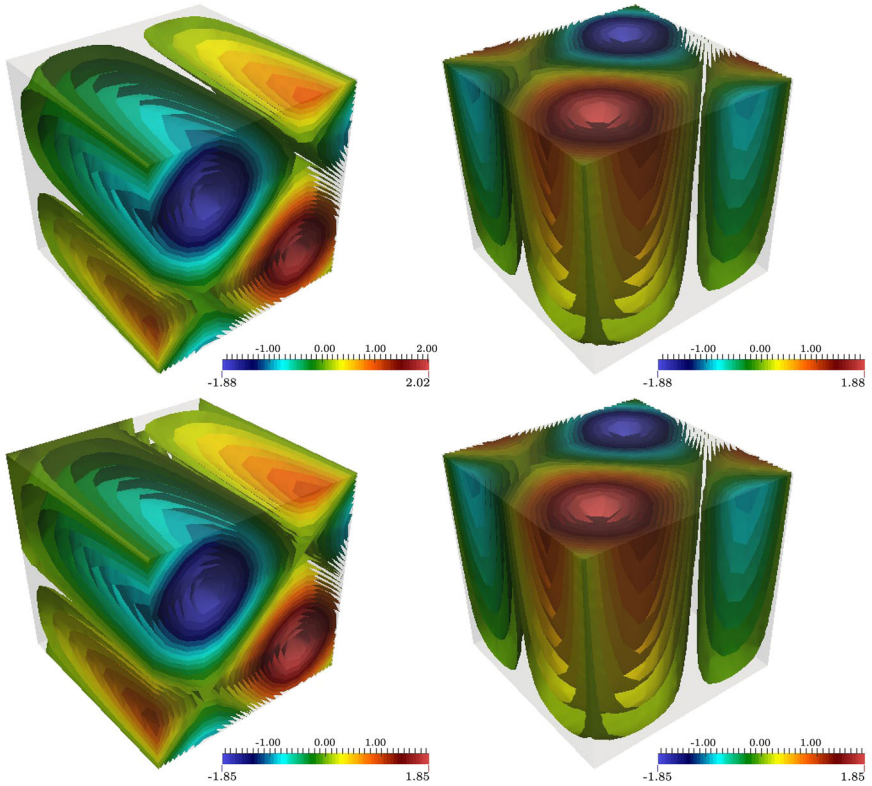

Fig. 7 Example 4, iso-surfaces of $u_{h,1}$ (*top-left*) and $u_{h,3}$ (*top-right*) for $k = 2$, and its exact values (*bottom*)

for all $\mathbf{x} := (x_1, x_2)^t \in \Omega$, where $\lambda := \frac{Re}{2} - \sqrt{\frac{Re^2}{4} + 4\pi^2}$ and $Re := \mu^{-1} = 10$ is the Reynolds number. It is easy to see in this linear case that $\alpha_0 = \gamma_0 = \mu$. Concerning the triangulations employed in our computations, we first consider seven meshes that are Cartesian refinements of a domain defined in terms of squares, and then we split each square into four congruent triangles.

In Example 2 we deal with the nonlinear version of Example 1. More precisely, we consider instead of $\mu = 0.1$ the kinematic viscosity function $\mu : R^+ \to R^+$ given by the Carreau law, that is $\mu(t) := \mu_0 + \mu_1(1+t^2)^{(\beta-2)/2} \quad \forall\, t \in R^+$, with $\mu_0 = \mu_1 = 0.5$ and $\beta = 1.5$. It is easy to check in this case that the assumptions (2.2) and (2.3) are satisfied with

$$\gamma_0 = \mu_0 + \mu_1 \left\{ \frac{|\beta - 2|}{2} + 1 \right\} \quad \text{and} \quad \alpha_0 = \mu_0 \,.$$

Then, we let again $\Omega := (-0.5, 1.5) \times (0, 2)$, and choose the data $\mathbf{f}$ and $\mathbf{g}$ so that the exact solution is the same from Example 1. The set of triangulations utilized is also as in Example 1.

Next, in Example 3 we use the same nonlinearity μ from Example 2, consider the L-shaped domain $\Omega := (-1, 1)^2 \setminus [0, 1]^2$, and choose the data $\mathbf{f}$ and $\mathbf{g}$ so that the exact solution is given by

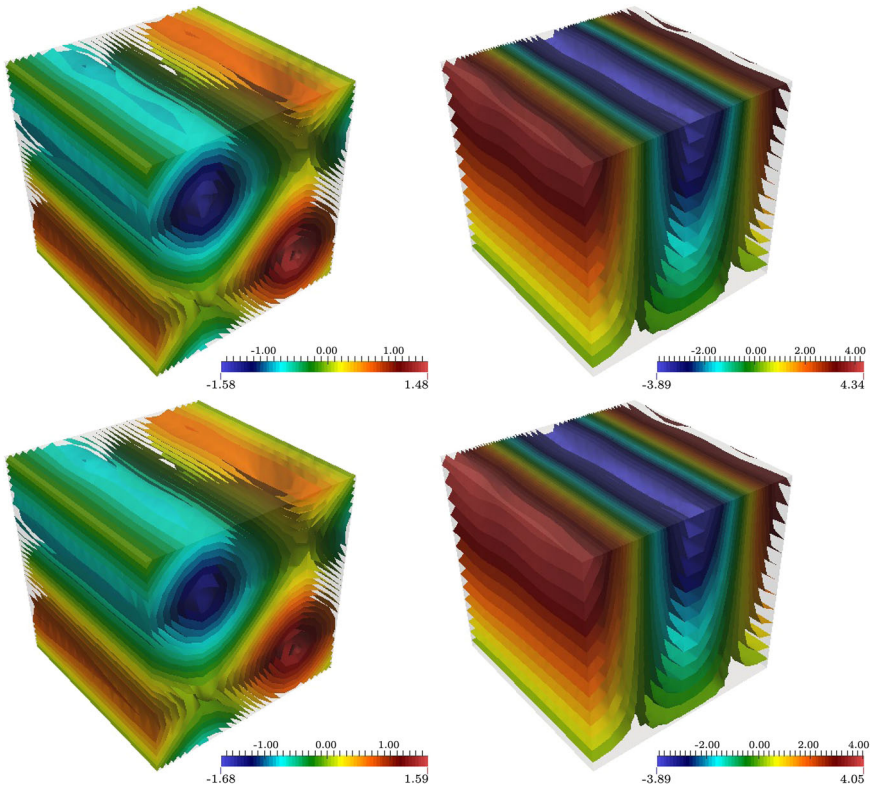

Fig. 8 Example 4, iso-surfaces of $\sigma_{h,11}$ (*top-left*) and $\sigma_{h,32}$ (*top-right*) for $k = 2$, and its exact values (*bottom*)

$$\mathbf{u}(\mathbf{x}) = \left(r^{2/3}\sin(\theta),\ -r^{2/3}\cos(\theta)\right),$$
$$p(\mathbf{x}) = \cos(x_1)\cos(x_2) - \sin^2(1),$$

for all $\mathbf{x} := (x_1, x_2)^{\mathsf{t}} \in \Omega$, where $r := |\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$ and $\theta := \arctan\left(\frac{x_2}{x_1}\right)$. We remark that $\nabla\mathbf{u}$ is singular at the origin, and hence lower rates of convergence are expected in our computations. The meshes are generated analogously to the previous examples.

Finally, in Example 4 we consider the three dimensional domain $\Omega := (0, 1)^3$, and assume the same kinematic viscosity function μ from Examples 2 and 3. In addition, the data $\mathbf{f}$ and $\mathbf{g}$ are chosen so that the exact solution is given by

$$\mathbf{u}(\mathbf{x}) = \Big(x_1(\sin(2\pi x_3) - \sin(2\pi x_2)),\ x_2(\sin(2\pi x_1) - \sin(2\pi x_3)),$$
$$x_3(\sin(2\pi x_2) - \sin(2\pi x_1))\Big),$$
$$p(\mathbf{x}) = x_1x_2x_3\sin(2\pi x_1)\sin(2\pi x_2)\sin(2\pi x_3) + \frac{1}{8\pi^3},$$

for all $\mathbf{x} := (x_1, x_2, x_3)^{\mathsf{t}} \in \Omega$.

It is easy to check that $\mathbf{u}$ is divergence free and $\int_\Omega p = 0$ for each one of the aforedescribed examples.

It is important to remark here that we do not provide any postprocessing for the velocity $\mathbf{u}$ in the numerical results shown below. Nevertheless, we can report that we did perform some preliminary numerical experiments by using the postprocessing formulas given in [12], which yielded exactly the same order of $\mathbf{u}_h$, that is $\mathcal{O}(h^k)$, and hence no superconvergence was observed. The alternative formula given in [11] will be considered in a forthcoming related work.

In Tables 1 and 2 we summarize the convergence history of the augmented HDG method (2.9) as applied to Examples 1 and 2 for the polynomial degrees $k \in \{1, 2, 3\}$. We observe there, looking at the experimental rates of convergence, that the orders predicted for each k by Theorems 4.1 and 4.2, and estimates (4.13) and (4.14), are attained by all the unknowns for these smooth examples. Actually, the errors $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Sigma_h}$ and $\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}$ behave exactly as proved, whereas the remaining ones show higher orders of convergence. In particular, $\|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \widehat{\mathbf{u}}_h\|_h$ presents a superconvergence phenomenon with two additional powers of h. In addition, it is interesting to notice that these numerical results provide the same rates of convergence obtained for the linear case in [11], and hence they might constitute numerical evidences supporting the conjecture that the a priori error estimates derived in the present paper are not sharp. We plan to address this issue in a separate work. Nevertheless, as already mentioned at the beginning of Sect. 4, whether the projection-based error analysis developed in [11] will work or not in this nonlinear case is still an open problem.

Furthermore, preliminary numerical experiments for Example 2, using degree k instead of $k-1$ in the definition of the subspace V_h, showed that the convergence rates are the same of Table 2. Perhaps, the only advantages of this modification with respect to the approach of the present paper are the possibility of using the polynomial degree $k = 0$ and the fact that the superconvergence behavior of the variable $\boldsymbol{\lambda}_h$ is recovered when $k = 1$. The above could very well mean that the restriction $k \geq 1$ and the degree $k - 1$ for defining V_h are just technical assumptions of our analysis. On the other hand, even though the estimates given in Sect. 4 hold for τ small enough, the results provided in Table 3 for Example 2 insinuate the robustness of our method within a larger, but still limited, range of variability of this parameter. Indeed, we observe there that for fixed values of k and h, the errors of some variables behave pretty much of the same order when larger values of τ (up to $\tau = 10$) are employed. However, while for even larger values of the parameter such as $\tau \in \{100, 1000\}$ the method does not break down, we notice that in this case some errors begin to increase.

In Table 4 we summarize the convergence history of the augmented HDG method (2.9) as applied to Example 3 for the polynomial degrees $k \in \{1, 2, 3\}$. In this case, and because of the singularity at the origin of the exact solution, the theoretical orders of convergence are far to be attained. In fact, similarly as obtained in [6], $\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}$ behaves as $\mathcal{O}(h^{\min\{k,4/3\}})$, whereas $\|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega} = \mathcal{O}(h^{2/3})$. Also, $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} = \mathcal{O}(h^{2/3})$, $\|\Pi_{\mathcal{E}_h}(\mathbf{u}) - \widehat{\mathbf{u}}_h\|_h = \mathcal{O}(h^{\min\{k,4/3\}})$, and thanks to (4.14), $\|p - p_h\|_{0,\Omega} = \mathcal{O}(h^{2/3})$ as well. Moreover, the behaviour of $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Sigma_h}$ is explained by the fact that the a priori estimate for $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Sigma_h}$ depends on the regularity of $\mathbf{div}(\boldsymbol{\sigma})$, which can be shown to belong precisely to $\mathbf{H}^{-1/3}(\Omega)$. A classical way of circumventing this drawback is the incorporation of an adaptive scheme based on a posteriori error estimates. This issue will also be addressed in a forthcoming paper.

On the other hand, in Table 5 we present the convergence history of the augmented HDG method (2.9) as applied to Example 4 for the polynomial degrees $k \in \{1, 2, 3\}$. The remarks in this case are exactly the same given above for Examples 1 and 2.

Finally, some components of the approximate and exact solutions for Examples 2, 3, and 4 are displayed in Figures 1, 2, 3, 4, 5, 6, 7 and 8. They all correspond to those obtained

with the fourth mesh and for the polynomial degree k indicated in each case. Here we use the notations $\mathbf{t}_h = (t_{h,ij})_{i,j=1,n}$, $\boldsymbol{\sigma}_h = (\sigma_{h,ij})_{i,j=1,n}$, and $\mathbf{u}_h = (u_{h,i})_{i=1,n}$.

Acknowledgments The authors are thankful to Paul Castillo and Manuel Solano for valuable remarks concerning the computational implementation of the HDG method.

References

1. Baranger, J., Najib, K., Sandri, D.: Numerical analysis of a three-fields model for a quasi-Newtonian flow. Comput. Methods Appl. Mech. Eng. **109**, 281–292 (1993)
2. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer, Berlin (1991)
3. Bustinza, R., Gatica, G.N.: A local discontinuous Galerkin method for nonlinear diffusion problems with mixed boundary conditions. SIAM J. Sci. Comput. **26**, 152–177 (2004)
4. Bustinza, R., Gatica, G.N.: A mixed local discontinuous Galerkin for a class of nonlinear problems in fluid mechanics. J. Comput. Phys. **207**, 427–456 (2005)
5. Carrero, J., Cockburn, B., Schötzau, D.: Hybridized, globally divergence-free LDG methods. Part I: the Stokes problem. Math. Comp. **75**, 533–563 (2006)
6. Castillo, P.E., Sequeira, F.A.: Computational aspects of the local discontinuous Galerkin method on unstructured grids in three dimensions. Math. Comput. Model. **57**, 2279–2288 (2013)
7. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Nort-Holland, Amsterdam (1978)
8. Clement, P.: Un Problème d'approximation Par éléments Finis, PhD thesis, Ecole Polytechnique Fédérale de Lausane (1973)
9. Cockburn, B., Dong, B., Guzmán, J.: A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. Math. Comp. **77**, 1887–1916 (2008)
10. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. **47**, 1319–1365 (2009)
11. Cockburn, B., Gopalakrishnan, J., Nguyen, N.C., Peraire, J., Sayas, F.J.: Analysis of HDG methods for Stokes flow. Math. Comput. **80**, 723–760 (2011)
12. Cockburn, B., Gopalakrishnan, J., Sayas, F.J.: A projection-based error analysis of HDG methods. Math. Comp. **79**, 1351–1367 (2010)
13. Cockburn, B., Guzmán, J., Wang, H.: Superconvergent discontinuous Galerkin methods for second-order elliptic problems. Math. Comp. **78**, 1–24 (2009)
14. Cockburn, B., Shi, K.: Devising HDG methods for Stokes flow: an overview. Comput. Fluids **98**, 221–229 (2014)
15. Cockburn, B., Shu, C.: The local discontinuous Galerkin method for time-dependent convection-diffusion systems. SIAM J. Numer. Anal. **35**, 2440–2463 (1998)
16. Congreve, S., Houston, P., Süli, E., Whiler, T.P.: Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems II: strongly monotone quasi-Newtonian flows. IMA J. Numer. Anal. **33**, 1386–1415 (2013)
17. Dubiner, M.: Spectral methods on triangles and other domains. J. Sci. Comput. **6**, 345–390 (1991)
18. Gatica, G.N.: A Simple Introduction to the Mixed Finite Element Method: Theory and Applications, SpringerBriefs in Mathematics. Springer, Berlin (2014)
19. Gatica, G.N., González, M., Meddahi, S.: A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. I: a priori error analysis. Comput. Methods Appl. Mech. Eng. **193**, 881–892 (2004)
20. Gatica, G.N., Heuer, N., Meddahi, S.: On the numerical analysis of nonlinear twofold saddle point problems. IMA J. Numer. Anal. **23**, 301–330 (2003)
21. Gatica, G.N., Márquez, A., Sánchez, M.A.: A priori and a posteriori error analyses of a velocity-pseudostress formulation for a class of quasi-Newtonian Stokes flows. Comput. Methods Appl. Mech. Eng. **200**, 1619–1636 (2011)
22. Girault, V., Raviart, P. A.: Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms, Springer Series in Computational Mathematics, vol. 5. Springer, Berlin (1986)
23. Hiptmair, R.: Finite elements in computational electromagnetism. Acta Numer. **11**, 237–339 (2002)
24. Houston, P., Robson, J., Süli, E.: Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems. I. The scalar case. IMA J. Numer. Anal. **25**, 726–749 (2005)
25. Howell, J.S.: Dual-mixed finite element approximation of Stokes and nonlinear Stokes problems using trace-free velocity gradients. J. Comput. Appl. Math. **231**, 780–792 (2009)

26. Kovasznay, L.I.G.: Laminar flow behind a two-dimensional grid. Proc. Camb. Philos. Soc. **44**, 58–62 (1948)
27. Ladyzhenskaya, O.: New equations for the description of the viscous incompressible fluids and solvability in the large for the boundary value problems of them. In: Boundary Value Problems of Mathematical Physics V, Providence, RI: AMS (1970)
28. Loula, A.F.D., Guerreiro, J.N.C.: Finite element analysis of nonlinear creeping flows. Comput. Methods Appl. Mech. Eng. **99**, 87–109 (1990)
29. Nguyen, N.C., Peraire, J., Cockburn, B.: An implicit high-order hybridizable discontinuous Galerkin method for nonlinear convection-diffusion equations. J. Comput. Phys. **228**, 8841–8855 (2009)
30. Nguyen, N.C., Peraire, J., Cockburn, B.: A hybridizable discontinuous Galerkin method for Stokes flow. Comput. Methods Appl. Mech. Eng. **199**, 582–597 (2010)
31. Roberts, J. E., Thomas, J. M.: Mixed and Hybrid Methods. In: Ciarlet, P.G., Lions, J.L. (eds). Handbook of Numerical Analysis, vol. II, Finite Element Methods (Part 1). Nort-Holland, Amsterdam (1991)
32. Sandri, D.: Sur l'approximation numérique des écoulements quasi-Newtoniens dont la viscosité suit la loi puissance ou la loi de Carreau. Math. Model. Numer. Anal. **27**, 131–155 (1993)
33. Schötzau, D., Schwab, C., Toselli, A.: Mixed *hp*-DGFEM for incompressible flows. SIAM J. Numer. Anal. **40**, 2171–2194 (2003)