SUPG-stabilized time-DG finite and virtual elements for the time-dependent advection–diffusion equation

L. Beirão da Veiga [a,b] , F. Dassi [a] , S. Gómez [a],*

[a] *Department of Mathematics and Applications, University of Milano-Bicocca, 20125 Milan, Italy*
[b] *IMATI-CNR, 27100 Pavia, Italy*

ARTICLE INFO

ABSTRACT

We carry out a stability and convergence analysis for the fully discrete scheme obtained by combining a finite or virtual element spatial discretization with the upwind-discontinuous Galerkin time-stepping applied to the time-dependent advection–diffusion equation. A space–time streamline-upwind Petrov–Galerkin term is used to stabilize the method. More precisely, we show that the method is inf–sup stable with constant independent of the diffusion coefficient, which ensures the robustness of the method in the convection- and diffusion-dominated regimes. Moreover, we prove optimal convergence rates in both regimes for the error in the energy norm. An important feature of the presented analysis is the control in the full $L^2(0, T; L^2(\Omega))$ norm without the need of introducing an artificial reaction term in the model. We finally present some numerical experiments in $(3 + 1)$-dimensions that validate our theoretical results.

1. Introduction

The present contribution focuses on the classical time-dependent advection–diffusion equations, also thought as a first step towards more complex nonlinear fluid dynamic problems. More specifically, let the space–time cylinder $Q_T = \Omega \times (0, T)$, where $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) is an open, bounded polytopic domain with Lipschitz boundary $\partial\Omega$, and let $T > 0$ represent the final time. Then, for given strictly positive diffusion coefficient ν, transport solenoidal field $\boldsymbol{\beta}$, source term f, and initial datum u_0, we consider the following advection–diffusion IBVP:

$$\begin{cases} \partial_t u - \nu \Delta u + \boldsymbol{\beta} \cdot \nabla u = f & \text{in } Q_T, \\ u = 0 & \text{on } \Gamma_{\mathrm{D}}, \\ u = u_0 & \text{on } \Sigma_0, \end{cases} \tag{1.1}$$

where the surfaces $\Sigma_0 := \Omega \times \{0\}$, $\Sigma_T := \Omega \times \{T\}$, and $\Gamma_{\mathrm{D}} := \partial\Omega \times (0, T)$.

In the numerical analysis literature, problem (1.1), in addition to its specific interest, has often represented an important step towards the study of more complex models, such as those describing incompressible fluid flows at high Reynolds numbers. This concurs in motivating the very large amount of articles dealing with the so called advection-dominated case, that is in the development and analysis of numerical methods able to deliver accurate and reliable solutions also when $|\boldsymbol{\beta}| \gg \nu$. Indeed, many stabilization techniques have been designed to address the well-known issue of spurious oscillations or instabilities of conforming

---

* Corresponding author.

E-mail addresses: lourenco.beirao@unimib.it (L. Beirão da Veiga), franco.dassi@unimib.it (F. Dassi), sergio.gomezmacias@unimib.it (S. Gómez).

finite element (FE) discretizations of model (1.1) in the advection-dominated regime. Such techniques include space–time least squares [1], streamline-upwind Petrov–Galerkin (SUPG) and variants [2–6], local projection stabilization (LPS) [7,8], and other symmetric stabilization terms [9]; see also the analysis, in an abstract framework, for spatial discretizations based on symmetric stabilization terms with discontinuous Galerkin time stepping in [10].

Some simplifications are commonly found in the literature, such as omitting the effect of the time discretizations [11], restricting to low-order time stepping schemes [5,9,12–14], and requiring a sufficiently strong reaction term σu such that (see, e.g., [7,12,15,16])

$$\inf_{(x,t) \in Q_T} \left(\sigma - \frac{1}{2} \operatorname{div} \boldsymbol{\beta} \right) \geq \sigma_0 > 0. \tag{1.2}$$

Note that the latter condition is relevant to higher order schemes in time, where the standard discrete energy argument only leads to control in $L^2(\Omega)$ at discrete time instants (plus the sum of the time-jumps at the time-mesh nodes) but not in $L^2(0, T; L^2(\Omega))$ and even less so in $L^\infty(0, T; L^2(\Omega))$. Condition (1.2) is typically justified by the fact that the problem for the variable $w = e^{-t/T} u$ satisfies such an assumption (at least when $\sigma - \frac{1}{2} \operatorname{div} \boldsymbol{\beta} \geq 0$). Although this is surely acceptable, we prefer to tackle the more complex (from the theoretical standpoint) case in which such transformation is not assumed so that no data modification is needed in the method, see Remark 3.8. This choice is also motivated by the possible extension to nonlinear problems, where the above transformation would induce the introduction of time-dependent factors also in front of the nonlinear terms. Similar observations hold for the more recent Virtual Element (VE) technology [17], its literature being clearly less rich than the FE one; some articles dealing with the above issue are [18–21].

The present work concerns the design and analysis of an SUPG-stabilized version of the fully discrete scheme obtained by combining a conforming FE or VE spatial discretization with an upwind-DG time stepping; in particular, we focus on the robustness analysis of the method in the advection-dominated regime (i.e., when $0 \leq \nu \ll 1$). Our main contributions are the following:

- We carry out the first stability and error analysis of a high-order-in-time SUPG-stabilized scheme for the time-dependent advection–diffusion IBVP (1.1), which does not require the transformation of the original problem, and does not rely on the presence of a positive reaction term. Although stability and optimal convergence of the upwind-discontinuous Galerkin (DG) SUPG-stabilized finite element method (FEM) have been hypothesized [22, §3.3], a thorough analysis was missing in the literature even for FEM (and less so for VEM, for which the proposed methodology is novel).
- We address, in a unified framework, conforming FE and VE spatial discretizations. Our analysis focuses on VE spaces; however, the same ideas apply to conforming FE spaces with some simplifications (as detailed in Section 3.4).
- We show an inf–sup estimate with stability constant independent of the meshsize, the time step, and the diffusion coefficient ν. Such an estimate is used to prove that, in a certain energy norm, the fully discrete solution satisfies: *i)* a continuous dependence on the data of the problem uniformly in ν, and *ii)* some *a priori* error bounds, which do not degenerate when the diffusion coefficient ν is small, and depend only on the interpolation and the nonconsistency errors.
- At the end of the article we evaluate the practical performance of the proposed scheme through a set of numerical tests in $(3 + 1)$-dimensions, for different orders of approximation in time and space.

The manuscript is organized as follows. In Section 2, we present some basic notation, the variational form of the continuous problem and a well-known stability result. In Section 3, we present the proposed VEM, discuss the particular case of FEM and the extension to Serendipity VEM, and present our main theoretical results. Section 4 is devoted to the proof of the well-posedness and stability of the scheme. In Section 5, we develop the convergence analysis, first addressing the general case of changing meshes (that is, when the spatial mesh can change form one time-slab to the next one) and then particularizing to the more favourable case with fixed spatial mesh. Finally, we present some numerical tests in Section 6 and make some concluding remarks in Section 7.

2. Basic notation and weak formulation of the problem.

We start by reviewing some basic notation we will use through the article. We denote the first-order time derivative operator by ∂_t, and the spatial gradient and Laplacian operators by ∇ and Δ, respectively. We will use standard notation for Sobolev spaces, seminorms, and norms [23]. For instance, given an open, bounded domain $Y \subset \mathbb{R}^d$ ($d \in \mathbb{N}$), and scalars $p \in [2, \infty]$ and $s \in \mathbb{R}$, we denote by $W^{s,p}(Y)$ the standard Sobolev space, and its associated seminorm and norm by $| \cdot |_{W^{s,p}(Y)}$ and $\| \cdot \|_{W^{s,p}(Y)}$, respectively. In particular, for $p = 2$, we use the notation $H^s(Y) = W^{s,2}(Y)$, and denote its associated seminorm and norm by $| \cdot |_{H^s(Y)}$ and $\| \cdot \|_{H^s(Y)}$, respectively. Moreover, the space $L^2(Y) = H^0(Y)$ denotes the space of Lebesgue square integrable functions over Y with its corresponding inner product $(\cdot, \cdot)_Y$, and $H_0^1(Y)$ denotes the space of functions in $H^1(Y)$ with zero trace on ∂Y. A superscript d is used to represent the seminorms and norms of vector fields with d-components. In addition, given a Banach space $(X, \| \cdot \|_X)$, a time interval (a, b), and a scalar $s \in \mathbb{R}$, we denote the Bochner–Sobolev space by $H^s(a, b; X)$. Finally, we use the following notation for the algebraic tensor product of two spaces, say V and W:

$$V \otimes W := \operatorname{span}\{vw \ : \ v \in V \text{ and } w \in W\}.$$

Given $k \in \mathbb{N}$, we denote the space of polynomials of degree at most k defined on Y by $\mathbb{P}_k(Y)$.

For the time being, we assume the following data regularity. The transport advective field $\boldsymbol{\beta} \in W^{1,\infty}(0, T; L^\infty(\Omega))$ with $\operatorname{div} \boldsymbol{\beta} = 0$, the source term $f \in L^2(Q_T)$, and the initial datum $u_0 \in L^2(\Omega)$. Then, denoting by $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ and $b : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$

the following bilinear forms:

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, d\boldsymbol{x} \quad \text{and} \quad b(u, v) = \int_\Omega v \boldsymbol{\beta} \cdot \nabla u \, d\boldsymbol{x}, \tag{2.1}$$

the continuous week formulation of the IBVP (1.1) reads (see [24, §7.1.1]): find $u \in L^2(0, T; H^1_0(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \subset C([0, T]; L^2(\Omega))$ such that $u = u_0$ on Σ_0, and for almost all $t \in (0, T)$, it holds

$$\langle \partial_t u, v \rangle + v a(u, v) + b(u, v) = (f, v)_\Omega \qquad \forall v \in H^1_0(\Omega), \tag{2.2}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality between $H^{-1}(\Omega)$ and $H^1_0(\Omega)$.

For any $\tilde{t} \in (0, T]$, integrating in time equation (2.2) over $(0, \tilde{t})$ and using the skew symmetry of the bilinear form $b(\cdot, \cdot)$ (i.e., $b(u, v) = -b(v, u)$), and the Hölder and the Young inequalities, we obtain the following bound:

$$\frac{1}{2} \|u(\cdot, \tilde{t})\|^2_{L^2(\Omega)} + v \|u\|^2_{L^2(0, \tilde{t}; H^1(\Omega))} \le \frac{1}{2} \|u_0\|^2_{L^2(\Omega)} + \|f\|^2_{L^1(0, \tilde{t}; L^2(\Omega))} + \frac{1}{4} \|u\|^2_{L^\infty(0, \tilde{t}; L^2(\Omega))}. \tag{2.3}$$

Since $u \in C^0([0, T]; L^2(\Omega))$, we can take the maximum over $[0, T]$ in (2.3) and deduce the following stability estimate:

$$\frac{1}{4} \|u\|^2_{L^\infty(0, T; L^2(\Omega))} + v \|u\|^2_{L^2(0, T; H^1(\Omega))} \le \frac{1}{2} \|u_0\|^2_{L^2(\Omega)} + \|f\|^2_{L^1(0, T; L^2(\Omega))}.$$

The above inequality shows a uniform-in-v continuous dependence of the solution to (2.2) on the data of the problem. Such a property is clearly desirable to be reproduced at the discrete level.

3. Description of the method and main results

In this section, we describe the proposed SUPG-stabilized time-DG VEM for the discretization of model (1.1) and present the major theoretical results. Some notation for tensor-product-in-time meshes is introduced in Section 3.1. In Section 3.2, we recall the definition of the local enhanced VE spaces in two and three dimensions, their corresponding degrees of freedom, and some computable polynomial projections. In Section 3.3, global discrete spaces are defined as the tensor product of the space of piecewise polynomials in time and $H^1_0(\Omega)$-conforming VE spaces, and we present the discrete bilinear forms in the definition of the SUPG-stabilized time-DG VEM in Section 3.5. Finally the main theoretical results, asserting the well posedness of the discrete problem and its convergence properties, are presented in Section 3.6.

3.1. Space–time mesh notation and assumptions

Let $\{\Omega_h\}_{h>0}$ be a family of polytopic partitions of the spatial domain $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ (see Remark 3.4 regarding the case with variable spatial meshes). For each $K \in \{\Omega_h\}_{h>0}$ and each facet F of K, we denote by h_K and h_F the diameters of K and F, respectively. We make the following assumption on the family $\{\Omega_h\}_{h>0}$.

Assumption 3.1 (*Mesh Regularity*). There exists a strictly positive constant ϱ such that the following conditions hold for any element $K \in \{\Omega_h\}_{h>0}$:

(A1) K is star-shaped with respect to a ball B_K of radius larger than or equal to ϱh_K;

(A2) (if $d = 2$) each facet F has length larger than or equal to ϱh_K;

(A3) (if $d = 3$) each facet F is star-shaped with respect to a disk B_F of radius larger than or equal to ϱh_K, and each edge e of the (polygonal) facet F has length larger than or equal to ϱh_K.

In particular, Assumptions (A1)–(A3) imply the existence of a uniform maximum number of facets for each element K of $\{\Omega_h\}_{h>0}$.

Let $\mathcal{T}_\tau$ be a partition of the time interval $(0, T)$ given by $0 := t_0 < t_1 < \cdots < t_N := T$. For $n = 1, \dots, N$, we define the time interval $I_n := (t_{n-1}, t_n)$, the surface $\Sigma_n := \Omega \times \{t_n\}$, and the time step $\tau_n := t_n - t_{n-1}$. We further define the spatial meshsize $h := \max_{K \in \Omega_h} h_K$, the minimum element diameter $h_{\min} := \min_{K \in \Omega_h} h_K$, and the maximum time step $\tau := \max_{n=1,\dots,N} \tau_n$.

Finally, for each $K \in \Omega_h$ and $n = 1, \dots, N$, we define the space–time prism $K_n := K \times I_n$.

Remark 3.2 (*Relaxation of the Mesh Assumptions*). The above regularity assumptions on the mesh could be partially relaxed, see for instance [25, §2.3], [26, §2 and §5.1], and [27, §3] for some results in this direction. ∎

3.2. Local virtual element spaces and projections

Let k and r be integer numbers such that $k \ge 1$ and $r \ge 0$, which denote the "degrees of approximation" in space and time, respectively.

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open $(d - j)$-polytope for some $j \in \{0, \dots, d - 1\}$. We denote by $\boldsymbol{x}_\mathcal{D}$ the centroid of $\mathcal{D}$, and introduce the following scaled and shifted monomial basis for the space $\mathbb{P}_k(\mathcal{D})$:

$$\mathcal{M}_k(\mathcal{D}) := \bigcup_{\ell=0}^k \mathbb{M}_\ell(\mathcal{D}) \quad \text{with} \quad \mathbb{M}_\ell(\mathcal{D}) := \left\{ m_{\boldsymbol{\alpha}} = \left(\frac{\boldsymbol{x} - \boldsymbol{x}_\mathcal{D}}{h_K} \right)^{\boldsymbol{\alpha}} \; : \; \boldsymbol{\alpha} \in \mathbb{N}^d \text{ with } |\boldsymbol{\alpha}| = \ell \right\}.$$

In the VE context, the use of these bases is particularly convenient for implementation.

We denote by $\Pi_k^{0,K} : L^2(K) \to \mathbb{P}_k(K)$ and $\boldsymbol{\Pi}_k^{0,K} : L^2(K)^d \to \mathbb{P}_k(K)^d$ the $L^2(K)$-orthogonal projection operators in $\mathbb{P}_k(K)$ and $\mathbb{P}_k(K)^d$, respectively. Moreover, we denote by $\Pi_k^{\nabla,K} : H^1(K) \to \mathbb{P}_k(K)$ the $H^1(K)$-orthogonal projection operator, defined for any $v \in H^1(K)$ as the solution to the following local problem:

$$\begin{cases} \displaystyle\int_K \nabla(\Pi_k^{\nabla,K} v - v) \cdot \nabla q_k \, \mathrm{d}\boldsymbol{x} = 0 & \forall q_k \in \mathbb{P}_k(K), \\ \displaystyle\int_{\partial K} (\Pi_k^{\nabla,K} - v) \, \mathrm{d}S = 0. \end{cases}$$

Two-dimensional virtual element spaces. If $\Omega \subset \mathbb{R}^2$, for each element $K \in \Omega_h$, we define the following local spaces:

$$\mathbb{B}_k(\partial K) := \{v \in C^0(\partial K) \, : \, v_{|e} \in \mathbb{P}_k(e) \text{ for each edge } e \text{ of } K\}, \tag{3.1a}$$

$$\widetilde{V}_k^{2D}(K) := \{v \in H^1(K) \, : \, v_{|\partial K} \in \mathbb{B}_k(\partial K) \text{ and } \Delta v \in \mathbb{P}_k(K)\}, \tag{3.1b}$$

$$V_k^{2D}(K) := \left\{v \in \widetilde{V}_k^{2D}(K) \, : \, \int_K (v - \Pi_k^{\nabla,K} v) m_{\boldsymbol{\alpha}} \, \mathrm{d}\boldsymbol{x} = 0 \text{ for all } m_{\boldsymbol{\alpha}} \in \mathbb{M}_{k-1}(K) \cup \mathbb{M}_k(K)\right\}, \tag{3.1c}$$

where the latter is the standard local enhanced VE space introduced in [28, §3]. The following linear functionals constitute a set of unisolvent degrees of freedom (DoFs) for $V_k^{2D}(K)$ (see [28, Prop. 2]):

(Dv1) the values of v at the vertices of K;

(Dv2) (if $k \geq 2$) the values of v at $(k-1)$ distinct internal points along each edge e of K;

(Dv3) (if $k \geq 2$) the following moments of v against the elements of $\mathcal{M}_{k-2}(K)$:

$$\frac{1}{|K|} \int_K v m_{\boldsymbol{\alpha}} \, \mathrm{d}\boldsymbol{x} \quad \forall m_{\boldsymbol{\alpha}} \in \mathcal{M}_{k-2}(K).$$

Three-dimensional virtual element spaces. If $\Omega \subset \mathbb{R}^3$, for each element $K \in \Omega_h$, we define the following local spaces:

$$\mathbb{W}_k(\partial K) := \{v \in C^0(\partial K) \, : \, v_{|F} \in V_k^{2D}(F) \text{ for each face } F \text{ of } K\}, \tag{3.2a}$$

$$\widetilde{V}_k^{3D}(K) := \{v \in H^1(K) \, : \, v_{|\partial K} \in \mathbb{W}_k(\partial K) \text{ and } \Delta v \in \mathbb{P}_k(K)\}, \tag{3.2b}$$

$$V_k^{3D}(K) := \left\{v \in \widetilde{V}_k^{3D}(K) \, : \, \int_K (v - \Pi_k^{\nabla,K} v) m_{\boldsymbol{\alpha}} \, \mathrm{d}\boldsymbol{x} = 0 \text{ for all } m_{\boldsymbol{\alpha}} \in \mathbb{M}_{k-1}(K) \cup \mathbb{M}_k(K)\right\}, \tag{3.2c}$$

where we have denoted by $V_k^{2D}(F)$ the local enhanced VE space on the face F, as F is contained in a (two-dimensional) plane. The following linear functionals constitute a set of unisolvent DoFs for $V_k^{3D}(K)$ (see [28, §4.1]):

(Dv1*) the values of v at the vertices of K;

(Dv2*) (if $k \geq 2$) the following moments of v on each edge e of K:

$$\frac{1}{|e|} \int_e v m_{\boldsymbol{\gamma}} \, \mathrm{d}s \qquad \forall m_{\boldsymbol{\gamma}} \in \mathcal{M}_{k-2}(e);$$

(Dv3*) (if $k \geq 2$) the following moments of v on each face F of K:

$$\frac{1}{|F|} \int_K v m_{\boldsymbol{\beta}} \, \mathrm{d}\boldsymbol{x} \qquad \forall m_{\boldsymbol{\beta}} \in \mathcal{M}_{k-2}(F);$$

(Dv4*) (if $k \geq 2$) the following moments of v on K:

$$\frac{1}{|K|} \int_K v m_{\boldsymbol{\alpha}} \, \mathrm{d}\boldsymbol{x} \qquad \forall m_{\boldsymbol{\alpha}} \in \mathcal{M}_{k-2}(K).$$

Henceforth, we will denote by $V_k(K)$ the local enhanced VE space, regardless of the spatial dimension d.

For any $K \in \Omega_h$ and $v \in V_k(K)$, the following polynomial projections are computable using the DoFs in sets **(Dv1)–(Dv3)** (if $d = 2$) or in sets **(Dv1*)–(Dv4*)** (if $d = 3$):

$$\Pi_k^{0,K} v, \quad \boldsymbol{\Pi}_k^{0,K} \nabla v, \quad \Pi_k^{\nabla,K} v.$$

3.3. Global space and bilinear forms

For $d = 2$ or $d = 3$, we define the global VE space

$$\mathcal{V}_h := \{v \in H_0^1(\Omega) \, : \, v_{|K} \in V_k(K) \, \forall K \in \Omega_h\}, \tag{3.3}$$

and the global space–time VE–DG space

$$\mathcal{V}_h^\tau := \prod_{n=1}^{N} \mathbb{P}_r(I_n) \otimes \mathcal{V}_h.$$

For any piecewise scalar function w and $m \in \{1, \dots, N\}$, we denote by $w^{(m)}$ the restriction of w to the time slab $\Omega \times I_m$. Moreover, for $n = 1, \dots, N-1$, we define the time jump ($[\![\cdot]\!]_n$) of w as follows:

$$[\![w]\!]_n(\boldsymbol{x}) := w(\boldsymbol{x}, t_n^-) - w(\boldsymbol{x}, t_n^+) \quad \forall \boldsymbol{x} \in \Omega,$$

where

$$w(\boldsymbol{x}, t_n^-) := \lim_{\varepsilon \to 0^+} w^{(n)}(\boldsymbol{x}, t_n - \varepsilon) \quad \text{and} \quad w(\boldsymbol{x}, t_n^+) := \lim_{\varepsilon \to 0^+} w^{(n+1)}(\boldsymbol{x}, t_n + \varepsilon).$$

We now introduce the discrete bilinear forms we use in the definition of the space–time VEM–DG formulation in Section 3.5 below. Henceforth, we denote by Id the identity operator. In the following, the projection operators defined in Section 3.2 are to be understood as applied pointwise in time.

The bilinear form $m_{h,\tau}^{\partial_t}(\cdot, \cdot)$. We define the upwind-DG VE discretization of the first-order time derivative operator $\partial_t(\cdot)$ as follows:

$$m_{h,\tau}^{\partial_t}(u_{h,\tau}, v_{h,\tau}) := \sum_{n=1}^{N} \int_{I_n} m_h\big(\partial_t u_{h,\tau}(\cdot, t), v_{h,\tau}(\cdot, t)\big) \, dt - \sum_{n=1}^{N-1} m_h\big([\![u_{h,\tau}]\!]_n, v_{h,\tau}(\cdot, t_n^+)\big) + m_h\big(u_{h,\tau}(\cdot, 0), v_{h,\tau}(\cdot, 0)\big),$$

where $m_h : \mathcal{V}_h \times \mathcal{V}_h \to \mathbb{R}$ is the standard VE discretization of the $L^2(\Omega)$-inner product, which can be written as

$$m_h(u_h, v_h) = \sum_{K \in \Omega_h} m_h^K(u_h, v_h),$$

with local contributions $m_h^K(\cdot, \cdot)$ given by

$$m_h^K(u_h, v_h) := (\Pi_k^{0,K} u_h, \Pi_k^{0,K} v_h)_K + s_m^K((\text{Id} - \Pi_k^{0,K})u_h, (\text{Id} - \Pi_k^{0,K})v_h),$$

for some symmetric bilinear form $s_m^K(\cdot, \cdot)$ chosen so that the following condition holds:

- **Stability of $s_m^K(\cdot, \cdot)$:** there exist positive constants $\check{\mu}$ and $\hat{\mu}$ independent of h and K, but depending on the degree k and the parameter ϱ in Assumption 3.1 such that

$$\check{\mu} \|v_h\|_{L^2(K)}^2 \le s_m^K(v_h, v_h) \le \hat{\mu} \|v_h\|_{L^2(K)}^2 \quad \forall v_h \in \ker(\Pi_k^{0,K}) \cap V_k(K). \tag{3.4}$$

Defining

$$\mu_* := \min\{1, \check{\mu}\} \quad \text{and} \quad \mu^* := \max\{1, \hat{\mu}\},$$

the stability property (3.4) implies

$$\mu_* \|v_h\|_{L^2(K)}^2 \le m_h^K(v_h, v_h) \le \mu^* \|v_h\|_{L^2(K)}^2 \quad \forall v_h \in V_k(K). \tag{3.5}$$

The bilinear form $a_{h,\tau}(\cdot, \cdot)$. We discretize the spatial Laplacian operator $(-\Delta)(\cdot)$ as follows:

$$a_{h,\tau}(u_{h,\tau}, v_{h,\tau}) := \sum_{n=1}^{N} \int_{I_n} a_h\big(u_{h,\tau}(\cdot, t), v_{h,\tau}(\cdot, t)\big) \, dt,$$

where $a_h : \mathcal{V}_h \times \mathcal{V}_h$ is the VE discretization of the bilinear form $a(\cdot, \cdot)$ in (2.1), which can be written as

$$a_h(u_h, v_h) = \sum_{K \in \Omega_h} a_h^K(u_h, v_h),$$

with local contributions $a_h^K(\cdot, \cdot)$ given by

$$a_h^K(u_h, v_h) := \big(\boldsymbol{\Pi}_{k-1}^{0,K} \nabla u_h, \boldsymbol{\Pi}_{k-1}^{0,K} \nabla v_h\big)_K + s_a^K\big((\text{Id} - \Pi_k^{\nabla,K})u_h, (\text{Id} - \Pi_k^{\nabla,K})v_h\big),$$

for some symmetric bilinear form $s_a^K(\cdot, \cdot)$ chosen so that the following condition holds:

- **Stability of $s_a^K(\cdot, \cdot)$:** there exist positive constants $\check{\alpha}$ and $\hat{\alpha}$ independent of h and K, but depending on the degree k and the parameter ϱ in Assumption 3.1 such that

$$\check{\alpha} \|\nabla v_h\|_{L^2(K)^d}^2 \le s_a^K(v_h, v_h) \le \hat{\alpha} \|\nabla v_h\|_{L^2(K)^d}^2 \quad \forall v_h \in \ker(\Pi_k^{\nabla,K}) \cap V_k(K). \tag{3.6}$$

Defining

$$\alpha_* := \min\{1, \check{\alpha}\} \quad \text{and} \quad \alpha^* := \max\{1, \hat{\alpha}\},$$

using the stability property (3.6), the stability of $\Pi_k^{0,K}$ in the $L^2(K)^d$-norm, and the fact that $\nabla \Pi_k^{\nabla,K} v_h \in \mathbb{P}_{k-1}(K)^d$, we deduce that

$$\alpha_* \|\nabla v_h\|_{L^2(K)^d}^2 \le a_h^K(v_h, v_h) \le \alpha^* \|\nabla v_h\|_{L^2(K)^d}^2 \qquad \forall v_h \in V_k(K). \tag{3.7}$$

The bilinear form $b_{h,\tau}^{\text{skew}}(\cdot,\cdot)$. As for the discretization of the advective term $(\beta \cdot \nabla u)$, we introduce the following skew-symmetric bilinear form:

$$b_{h,\tau}^{\text{skew}}(u_{h,\tau}, v_{h,\tau}) := \frac{1}{2} \sum_{n=1}^{N} \int_{I_n} \Big(b_h\big(u_{h,\tau}(\cdot,t), v_{h,\tau}(\cdot,t)\big) - b_h\big(v_{h,\tau}(\cdot,t), u_{h,\tau}(\cdot,t)\big) \Big) \, dt,$$

where $b_h : \mathcal{V}_h \times \mathcal{V}_h \to \mathbb{R}$ is given by

$$b_h(u_h, v_h) := \sum_{K \in \Omega_h} \big(\beta \cdot \Pi_k^{0,K} \nabla u_h, \, \Pi_k^{0,K} v_h\big)_K. \tag{3.8}$$

The above form (3.8) corresponds to the choice in [19, Eq. (4.5)] (see also [18, Eq. (15)]); in the two-dimensional case, it can be substituted with the alternative choice in [19, Eq. (4.6)].

The bilinear form $s_{h,\tau}^{\text{supg}}(\cdot,\cdot)$. Finally, we introduce the SUPG-stability bilinear form

$$s_{h,\tau}^{\text{supg}}(u_{h,\tau}, v_{h,\tau}) := \sum_{n=1}^{N} \sum_{K \in \Omega_h} s_{h,\tau}^{K_n,\text{supg}}(u_{h,\tau}, v_{h,\tau}),$$

with

$$s_{h,\tau}^{K_n,\text{supg}}(u_{h,\tau}, v_{h,\tau}) := \lambda_{K_n} \big(\mathcal{L}_{h,\tau}^K u_{h,\tau}, \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau}\big)_{K_n} + \beta_{K_n}^2 \lambda_{K_n} \int_{I_n} s_a^K\big((\text{Id} - \Pi_k^{\nabla,K})u_{h,\tau}, (\text{Id} - \Pi_k^{\nabla,K})v_{h,\tau}\big) \, dt,$$

for some parameter $\lambda_{K_n} > 0$ to be specified later, $\beta_{K_n} := \max\{\beta_\varepsilon, \|\beta\|_{L^\infty(K_n)^d}\}$ for some mesh-independent strictly positive "safeguard" constant β_ε, the stability term $s_a^K(\cdot,\cdot)$ as in the definition of $a_h(\cdot,\cdot)$, and the linear operators $\mathcal{L}_{h,\tau}^K$ and $\widetilde{\mathcal{L}}_{h,\tau}^K$ defined as follows:

$$\widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau} := \partial_t \Pi_k^{0,K} v_{h,\tau} + \beta \cdot \Pi_{k-1}^{0,K} \nabla v_{h,\tau}, \tag{3.9a}$$

$$\mathcal{L}_{h,\tau}^K u_{h,\tau} := \partial_t \Pi_k^{0,K} u_{h,\tau} - \nu \operatorname{div} \Pi_{k-1}^{0,K} \nabla u_{h,\tau} + \beta \cdot \Pi_{k-1}^{0,K} \nabla u_{h,\tau}. \tag{3.9b}$$

For convenience, we also define

$$\bar{\beta}_{Q_T} := \|\beta\|_{L^\infty(Q_T)^d}.$$

and assume, up to suitable scalings of the data, that $\bar{\beta}_{Q_T} \simeq 1$.

Remark 3.3 (*Stability Terms*). There exists a very large literature concerning different choices for the stabilization terms for VE discretizations and developing the associated theoretical support. Explicit expressions for the definition of the stability terms $s_m^K(\cdot,\cdot)$ and $s_a^K(\cdot,\cdot)$ can be found, for instance in [17,28,29] and some related proofs for instance in [25,26]; see also [30] for a recent discussion on the role of the stability terms for virtual element methods. ∎

3.4. Finite element and serendipity VE spaces

The proposed method immediately extends to the case of the classical Lagrangian FEM. If the mesh is simplicial, one can substitute the local spaces (3.1c) (and (3.2c) in three space dimensions) with

$$V_k(K) := \mathbb{P}_k(K) \qquad \forall K \in \Omega_h,$$

thus obtaining, c.f. (3.3), the standard Lagrangian FE space. In such a case, the scheme boils down to a standard SUPG-stabilized time-DG FEM approach, as all the polynomial projections appearing in the definition of the discrete forms disappear and, for the same reason, the stability terms vanish. Therefore, the theoretical results of this article trivially extend to such (simpler) case, yielding new results also for classical FEMs. Indeed, on our knowledge, results of this kind are missing in the literature, as previous SUPG schemes are low-order accurate in time (see, e.g., [5,12–14]) and the analysis of many techniques rely on the presence of a reaction term (see, e.g., [7,15], the discussion in the introduction of this contribution and Remark 3.8).

Another variant that can be considered is that of Serendipity Virtual Elements, which is a construction allowing to reduce the number of DoFs, an asset which is particularly useful for high-order approaches as the present one. We refer to [31] for a detailed presentation of Serendipity VEM (see also [32] for the associated interpolation and stability analysis) and here limit ourselves to a very brief review of the construction in three space dimensions. The idea is to eliminate DoFs that are internal to faces (since those that are internal to elements can be statically condensed) by introducing, for every face F of the polyhedral mesh, a projection operator

$$\Pi_F^{\mathsf{S}} : \widetilde{V}_k^{\text{2D}}(F) \longrightarrow \mathbb{P}_k(F)$$

that depends only on the DoFs associated with the boundary of F (vertex values and edge pointwise values). Clearly, such an operator can be constructed only if the $\mathbb{P}_k$-bubbles space on F reduces to $\{0\}$, a condition that depends on the geometry of F and

on the polynomial degree k. Alternatively, one can use an extended construction (see [31, §3]), but in the present brief review we prefer to stick to the simpler case where no such bubbles exist. Once such an operator is available, one can introduce the smaller space

$$V_k^{S,2D}(F) := \big\{ v \in \widetilde{V}_k^{2D}(F) \ : \ \int_F (v - \Pi_F^S v) m_\alpha \, d\boldsymbol{x} = 0 \ \text{ for all } m_\alpha \in \mathbb{M}_k(F) \big\},$$

whose associated DoFs are only **(Dv1)** and **(Dv2)**. Afterwards, one follows the same identical **3D** construction as in Section 3.2 but substituting the face spaces in (3.2a) with its Serendipity variant

$$\mathbb{W}_k^S(\partial K) := \{ v \in C^0(\partial K) \ : \ v_{|f} \in V_k^{S,2D}(f) \text{ for any face } f \text{ of } K \},$$

and using such a boundary space in (3.2b) instead of $\mathbb{W}_k(\partial K)$.

The final space has only degrees of freedom of type **(Dv1*)**, **(Dv2*)**, and **(Dv4*)** but retains all the key properties, such as approximation, of the original VE space [31,32].

3.5. SUPG-stabilized time-DG VEM

The proposed SUPG-stabilized time-DG VEM variational formulation is: find $u_{h,\tau} \in \mathcal{V}_h^\tau$ such that

$$\mathcal{B}_{h,\tau}^{\text{supg}}(u_{h,\tau}, v_{h,\tau}) = \ell_{h,\tau}^{\text{supg}}(v_{h,\tau}) \quad \forall v_{h,\tau} \in \mathcal{V}_h^\tau, \tag{3.10}$$

where

$$\mathcal{B}_{h,\tau}^{\text{supg}}(u_{h,\tau}, v_{h,\tau}) := m_{h,\tau}^{\partial_t}(u_{h,\tau}, v_{h,\tau}) + \nu a_{h,\tau}(u_{h,\tau}, v_{h,\tau}) + b_{h,\tau}^{\text{skew}}(u_{h,\tau}, v_{h,\tau}) + s_{h,\tau}^{\text{supg}}(u_{h,\tau}, v_{h,\tau}), \tag{3.11}$$

and

$$\ell_{h,\tau}^{\text{supg}}(v_{h,\tau}) := \sum_{K \in \Omega_h} \sum_{n=1}^N \Big(f, \ \Pi_k^{0,K} v_{h,\tau} + \lambda_{K_n} \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau} \Big)_{K_n} + \sum_{K \in \Omega_h} \big(u_0, \ \Pi_k^{0,K} v_{h,\tau}(\cdot, 0) \big)_K. \tag{3.12}$$

Remark 3.4. The method proposed above, and (unless clearly stated as will happen in Section 5.3) the stability and convergence analysis here developed apply identically to the case when the spatial mesh Ω_h changes at every time slab, that is we have a different mesh Ω_h^n for all $n = 0, 1, \ldots, N$. Nevertheless, in order to allow for a simpler notation and a clearer exposition, we prefer to keep the above simpler setting in the following developments. ∎

3.6. Main theoretical results

In this section, we present the main theoretical results of the article, namely the well-posedness of the discrete problem (3.10) and the associated quasi-robust error bounds. The proofs are postponed to Sections 4 and 5, respectively.

In what follows, we write $a \lesssim b$ to indicate the existence of a positive constant C independent of the meshsize h, the time step τ, and the diffusion coefficient ν such that $a \leq Cb$. Moreover, we write $a \simeq b$ meaning that $a \lesssim b$ and $b \lesssim a$.

We need to introduce some preliminary quantities. Let the upwind-jump functional

$$|u_{h,\tau}|_j^2 := \frac{1}{2} \Big(\|u_{h,\tau}\|_{L^2(\Sigma_T)}^2 + \sum_{n=1}^{N-1} \|[\![u_{h,\tau}]\!]_n\|_{L^2(\Omega)}^2 + \|u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \Big), \tag{3.13}$$

and the SUPG-functional

$$|u_{h,\tau}|_{\text{supg}}^2 := \sum_{n=1}^N \sum_{K \in \Omega_h} |u_{h,\tau}|_{K_n,\text{supg}}^2,$$

where

$$|u_{h,\tau}|_{K_n,\text{supg}}^2 := \lambda_{K_n} \|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + \lambda_{K_n} \beta_{K_n}^2 \|\nabla(\text{Id} - \Pi_k^{\nabla,K}) u_{h,\tau}\|_{L^2(K_n)^d}^2. \tag{3.14}$$

We also define the following norm in $\mathcal{V}_h^\tau$:

$$\|\|u_{h,\tau}\|\|_{\mathcal{E}}^2 := \|u_{h,\tau}\|_{L^2(Q_T)}^2 + |u_{h,\tau}|_j^2 + \nu \|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2 + |u_{h,\tau}|_{\text{supg}}^2. \tag{3.15}$$

The following assumption will be adopted in the sequel.

Assumption 3.5. For any $K \in \Omega_h$ and $n = 1, \ldots, N$, let λ_{K_n} be chosen so that

$$\lambda_{K_n} \leq \zeta \min\Big\{ \frac{h_K^2}{\nu C_{\text{inv}}^2}, \frac{h_K}{\bar{\beta}_{Q_T}} \Big\}, \tag{3.16}$$

with C_{inv} an inverse-estimate constant (c.f. Lemma 4.1 below) and some scalar constant ζ independent of h, τ, and ν. Moreover, let the following mild condition hold for some positive constant C_* independent of h and ν:

$$\tau \leq C_* h_{\min}^{\frac{1}{2}}. \tag{3.17}$$

The proposed method satisfies the following inf–sup estimate with stability constant independent of the diffusion coefficient ν; as a consequence, method (3.10) is well posed and remains stable even in the advection-dominated regime ($0 < \nu \ll 1$). The proof can be found in Section 4.

Theorem 3.6 (*Inf-Sup Stability*). *There exist positive constants C_*, ζ, and γ_I independent of h, τ, and ν such that, if $\tau \leq C_* h_{\min}^{\frac{1}{2}}$, and the stability parameters λ_{K_n} are chosen so that (3.16) is satisfied, it holds*

$$\||u_{h,\tau}\||_{\mathcal{E}} \leq \gamma_I \sup_{v_{h,\tau} \in \mathcal{V}_h^\tau \setminus \{0\}} \frac{\mathcal{B}_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, v_{h,\tau})}{\||v_{h,\tau}\||_{\mathcal{E}}} \quad \forall v_{h,\tau} \in \mathcal{V}_h^\tau.$$

The well-posedness of the scheme is a simple consequence of the above result (see the end of Section 4) and is stated here below.

Corollary 3.7 (*Well-Posedness*). *Under the assumptions of Theorem 3.6, there exists a unique solution $u_{h,\tau} \in \mathcal{V}_h^\tau$ to the discrete variational formulation (3.10). Moreover, if h and ζ are such that $\lambda_{K_n} \leq 1$ for all $K \in \Omega_h$ and $n \in \{1, \ldots, N\}$, the following continuous dependence on the data is satisfied:*

$$\||u_{h,\tau}\||_{\mathcal{E}} \leq \sqrt{2}\gamma_I \left(\|u_0\|_{L^2(\Omega)} + \|f\|_{L^2(Q_T)} \right).$$

The constant 1 appearing above in the condition $\lambda_{K_n} \leq 1$ was introduced only for simplicity of exposition. What is actually required for the well-posedness is that all the λ_{K_n} are uniformly bounded (in practice the λ_{K_n} are expected to be "small", c.f. (3.16)).

Remark 3.8 (*Stability in $L^2(0, T; L^2(\Omega))$*). Theorem 3.6 shows the stability of the fully discrete scheme in a norm including $L^2(0, T; L^2(\Omega))$ without resorting to the exponential transformation $w = e^{-t/T} u$, thus underlining that also the method applied directly to (1.1) is able to recover the full stability properties of the continuous problem. ∎

We now introduce the main *a priori* error estimates for our method, yielding optimal convergence rates in the energy norm $\||\cdot\||_{\mathcal{E}}$ defined in (3.15). The error estimates hold under the following assumption on the regularity of the data and the exact solution.

Assumption 3.9 (*Data Assumption*). For all $K \in \Omega_h$ and $n \in \{1, \ldots, N\}$, the solution u to the continuous weak formulation in (2.2), the source term f, and the initial condition u_0 satisfy:

$$u \in H^{q+1}(I_n; H^1(K)) \cap H^1(I_n; H^{s+1}(K)), \ f \in L^2(I_n; H^{s+1}(K)), \ \text{and} \ u_0 \in H^{s+1}(K),$$

and the advective field β satisfies,

$$\beta \in L^\infty(0, T; W^{s+1,\infty}(\Omega)) \cap W^{1,\infty}(0, T; L^\infty(\Omega)),$$

for $0 \leq q \leq r$ and $0 \leq s \leq k$.

The following convergence result holds (the proof can be found in Section 5).

Theorem 3.10 (*A Priori Error Estimates*). *Let Assumption 3.1 (on the mesh-regularity) and Assumption 3.5 (on the choice of λ_{K_n} and the relation of τ and $h_{\min}$) hold. Let also Assumption 3.9 hold with $q = s = r = k$. If the time steps $\{\tau_n\}_{n=1}^N$ are quasi-uniform and the solution u is sufficiently regular, we have the following error estimates (all bounds being uniform in ν):*

- Advection-dominated regime ($\nu \ll h_K$): $\lambda_{K_n} = \zeta \frac{h_K}{\beta_{Q_T}} \simeq h_K$

$$\||u - u_{h,\tau}\||_{\mathcal{E}}^2 \lesssim \tau^{2k+1} + h^{2k+1} + h^2 \tau^{2k} + h^{2k+1}\left(h/\tau + h^2/\tau^2\right) + \tau^{2k+2}/h_{\min}. \tag{3.18}$$

- Diffusion-dominated regime ($\nu \simeq 1$, $h_K \lesssim \nu$): $\lambda_{K_n} = \zeta \frac{h_K^2}{\nu C_{\mathrm{inv}}^2} \simeq h_K^2$

$$\||u - u_{h,\tau}\||_{\mathcal{E}}^2 \lesssim \tau^{2k+1} + h^{2k} + h^2 \tau^{2k} + h^{2k+2}/\tau^2 + \tau^{2k+2}/h_{\min}^2. \tag{3.19}$$

The quantities with negative powers of $h_{\min}$ in (3.18)–(3.19) suggest assuming quasi-uniformity also of the spatial mesh. In such a case, whenever the meshsize h and the time step τ are orthotropic (i.e., $\tau \simeq h$) we immediately observe that the error satisfies $\||u - u_{h,\tau}\||_{\mathcal{E}} \lesssim h^k$ in diffusion dominated cases and $\||u - u_{h,\tau}\||_{\mathcal{E}} \lesssim h^{k+1/2}$ in advection dominated cases, which are the optimal behaviours expected for quasi-robust schemes. Another important observation is that, since reducing τ is computationally cheaper than reducing h, one may be interested in the case $\tau \ll h$, representing the situation in which the time mesh is substantially finer than the spatial mesh. The presence of terms of the kind h/τ in the error estimate above are detrimental in this respect: we investigate the possibility of eliminating such terms in Section 5.3.

4. Well-posedness of the method

This section is devoted to prove Theorem 3.6 and Corollary 3.7.

4.1. Some useful tools

In the proof of the inf–sup stability estimate in Theorem 3.6 we make use of the following auxiliary exponential weight function:

$$\varphi = \varphi(t) := T \exp((T - t)/T), \tag{4.1}$$

which satisfies the following two important uniform bounds:

$$T \leq \varphi(t) \leq eT \qquad \forall t \in [0, T], \tag{4.2a}$$

$$-1 \leq \varphi'(t) \leq -e^{-1} \quad \forall t \in [0, T]. \tag{4.2b}$$

We denote by $\Pi_r^t : L^2(0, T) \to \mathbb{P}_k(\mathcal{T}_\tau)$ the $L^2(0, T)$-orthogonal projection operator in $\mathbb{P}_r(\mathcal{T}_\tau)$. In what follows, the operator Π_r^t is to be understood as applied pointwise in space.

We start by some inverse estimates for VE functions and polynomials.

Lemma 4.1 (*Local Inverse Estimates*). *Let Ω_h satisfy Assumption 3.1. Then, for all $K \in \Omega_h$ and $n = 1, \dots, N$, the following bounds hold:*

$$\|\nabla w_{r,k}\|_{L^2(I_n; L^2(K)^d)} \leq C_{\text{inv}} h_K^{-1} \|w_{r,k}\|_{L^2(I_n; L^2(K))} \qquad \forall w_{r,k} \in \mathbb{P}_r(I_n) \otimes V_k(K), \tag{4.3a}$$

$$\|\partial_t w_r\|_{L^2(I_n; L^2(K))} \leq C_{\text{inv}} \tau_n^{-1} \|w_r\|_{L^2(I_n; L^2(K))} \qquad \forall w_r \in \mathbb{P}_r(I_n) \otimes L^2(K), \tag{4.3b}$$

$$\|\operatorname{div} \mathbf{p}_{r,k}\|_{L^2(I_n; L^2(K))} \leq C_{\text{inv}} h_K^{-1} \|\mathbf{p}_{r,k}\|_{L^2(I_n; L^2(K)^d)} \qquad \forall \mathbf{p}_{r,k} \in \mathbb{P}_r(I_n) \otimes \mathbb{P}_k(K)^d, \tag{4.3c}$$

for some positive constant C_{inv} independent of K, h_K, and τ_n.

Proof. The inverse estimate in (4.3a) follows from the tensor-product structure of the element $K \times I_n$ and the space $\mathbb{P}_r(I_n) \otimes V_k(K)$, and the inverse estimates for VE functions in [33, Lemma 6.1] (for $d = 2$) and [34, Thm. 3.4] (for $d = 3$). The inverse estimates (4.3b) and (4.3c) are standard; see e.g., [35, §4.5]. □

In next lemma, we recall some approximation properties of Π_r^t from [36, Lemma 4.3].

Lemma 4.2. *There exists a positive constant C_S independent of h and τ such that, for $n = 1, \dots, N$ and $K \in \Omega_h$, the following bounds hold:*

$$\|(\operatorname{Id} - \Pi_r^t)(\varphi w_\tau)\|_{L^2(K_n)} \leq C_S \tau_n \|w_\tau\|_{L^2(K_n)} \qquad \forall w_\tau \in \mathbb{P}_r(I_n) \otimes L^2(K), \tag{4.4a}$$

$$\|\partial_t(\operatorname{Id} - \Pi_r^t)(\varphi w_\tau)\|_{L^2(K_n)} \leq C_S \|w_\tau\|_{L^2(K_n)} \qquad \forall w_\tau \in \mathbb{P}_r(I_n) \otimes L^2(K), \tag{4.4b}$$

$$\|(\operatorname{Id} - \Pi_r^t)(\varphi w_\tau)\|_{L^2(\Sigma_{n-1})} \leq C_S \tau_n^{\frac{1}{2}} \|w_\tau\|_{L^2(I_n; L^2(\Omega))} \qquad \forall w_\tau \in \mathbb{P}_r(I_n) \otimes L^2(\Omega), \tag{4.4c}$$

$$\|(\operatorname{Id} - \Pi_r^t)(\varphi w_\tau)\|_{L^2(\Sigma_n)} \leq C_S \tau_n^{\frac{1}{2}} \|w_\tau\|_{L^2(I_n; L^2(\Omega))} \qquad \forall w_\tau \in \mathbb{P}_r(I_n) \otimes L^2(\Omega). \tag{4.4d}$$

Proof. We show only the proof of (4.4b); the other bounds can be derived with very similar arguments. We start by introducing $\overline{\varphi} := \Pi_0^t \varphi$ and applying some simple steps

$$\|\partial_t(\operatorname{Id} - \Pi_r^t)(\varphi w_\tau)\|_{L^2(K_n)} = \|\partial_t(\operatorname{Id} - \Pi_r^t)((\varphi - \overline{\varphi}) w_\tau)\|_{L^2(K_n)} \leq \|\partial_t((\varphi - \overline{\varphi}) w_\tau)\|_{L^2(K_n)}$$
$$+ \|\partial_t \Pi_r^t((\varphi - \overline{\varphi}) w_\tau)\|_{L^2(K_n)} \leq \|\varphi' w_\tau\|_{L^2(K_n)} + \|(\varphi - \overline{\varphi}) \partial_t w_\tau\|_{L^2(K_n)} + \|\partial_t \Pi_r^t((\varphi - \overline{\varphi}) w_\tau)\|_{L^2(K_n)}. \tag{4.5}$$

The first term on the right-hand side is bounded trivially by $\|w_\tau\|_{L^2(K_n)}$ using (4.2b). The second term is bounded first by standard approximation properties of constant polynomials, afterwards by recalling (4.2b) and (4.3b):

$$\|(\varphi - \overline{\varphi}) \partial_t w_\tau\|_{L^2(K_n)} \leq \|\varphi - \overline{\varphi}\|_{L^\infty(I_n)} \|\partial_t w_\tau\|_{L^2(K_n)} \leq C \tau_n \|\varphi'\|_{L^\infty(I_n)} \|\partial_t w_\tau\|_{L^2(K_n)} \leq C \|w_\tau\|_{L^2(K_n)},$$

where C is a generic constant independent of τ_n.

The last term in (4.5) is bounded similarly. Since the function is polynomial in time we can apply an inverse estimate, afterwards use the continuity of the projection operator, and finally deploy again standard approximation properties of constant polynomials. We obtain

$$\|\partial_t \Pi_r^t((\varphi - \overline{\varphi}) w_\tau)\|_{L^2(K_n)} \leq C \tau_n^{-1} \|\Pi_r^t((\varphi - \overline{\varphi}) w_\tau)\|_{L^2(K_n)} \leq C \tau_n^{-1} \|(\varphi - \overline{\varphi}) w_\tau\|_{L^2(K_n)} \leq C \|w_\tau\|_{L^2(K_n)},$$

which completes the proof. □

Furthermore, the following estimate follows from [37, Lemma A.1(e)] and the equivalence of Π_0^t and the averaged Taylor polynomial T^0 defined in [37, Eq. (A.1)]: for all $K \in \Omega_h$ and $n = 1, \ldots, N$, it holds

$$\|\boldsymbol{\beta} - \Pi_0^t \boldsymbol{\beta}\|_{L^\infty(I_n; L^q(K))} \leq C_0 \tau_n \|\partial_t \boldsymbol{\beta}\|_{L^\infty(I_n; L^q(K))} \quad \forall q \in [1, \infty], \tag{4.6}$$

for some positive constant C_0 independent of h and τ.

We finally prove a simple orthogonality property for generic bilinear forms $s(\cdot, \cdot)$ on $V_k(K)$.

Lemma 4.3. *Let φ be defined as in* (4.1) *and $s(\cdot, \cdot)$ be a bilinear form on $V_k(K)$. For any $K \in \Omega_h$ and $n = 1, \ldots, N$, it holds*

$$\int_{I_n} s(u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi v_{h,\tau})) \, dt = 0 \qquad \forall u_{h,\tau}, v_{h,\tau} \in \mathbb{P}_r(I_n) \otimes V_k(K). \tag{4.7a}$$

Proof. Let $K \in \Omega_h$, $n \in \{1, \ldots, N\}$, and $u_{h,\tau}, v_{h,\tau} \in \mathbb{P}_r(I_n) \otimes V_k(K)$. Moreover, let $\{\phi_\ell\}_{\ell=1}^{\dim(V_k(K))}$ be a basis for the space $V_k(K)$. Then, there exist polynomials $\{\alpha_\ell\}_{\ell=1}^{\dim(V_k(K))} \subset \mathbb{P}_r(I_n)$ and $\{\beta_m\}_{m=1}^{\dim(V_k(K))} \subset \mathbb{P}_r(I_n)$ such that

$$u_{h,\tau}(x,t) = \sum_{\ell=1}^{\dim(V_k(K))} \alpha_\ell(t) \phi_\ell(x) \qquad\qquad \forall (x,t) \in K_n,$$

$$(\mathrm{Id} - \Pi_r^t)(\varphi v_{h,\tau})(x,t) = \sum_{m=1}^{\dim(V_k(K))} (\mathrm{Id} - \Pi_r^t)\big(\varphi(t)\beta_m(t)\big) \phi_m(x) \qquad\qquad \forall (x,t) \in K_n.$$

Therefore,

$$\int_{I_n} s(u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi v_{h,\tau})) \, dt = \sum_{\ell=1}^{\dim(V_k(K))} \sum_{m=1}^{\dim(V_k(K))} s(\phi_\ell, \phi_m) \int_{I_n} \alpha_\ell(t)(\mathrm{Id} - \Pi_r^t)(\varphi(t)\beta_m(t)) \, dt = 0,$$

which completes the proof. □

4.2. Inf–sup stability

For the sake of clarity, in next Lemmas we show some bounds that will be used to prove the inf–sup stability estimate in Theorem 3.6.

Lemma 4.4. *Let Assumption 3.5 hold. Then, under the same notation, for any $u_{h,\tau} \in \mathcal{V}_h^\tau$ and $v_{h,\tau} := \Pi_r^t(\varphi u_{h,\tau}) + \theta u_{h,\tau} \in \mathcal{V}_h^\tau$, with φ defined in* (4.1) *and some positive constant θ, the following bound holds:*

$$\|\|v_{h,\tau}\|\|_{\mathcal{E}} \leq \sqrt{2} \Big[2(eT)^2 + 2C_S^2 \tau + 4\zeta \big((1 + C_S^2)\bar{\beta}_{Q_T}^{-1} h + \delta^2 \bar{\beta}_{Q_T} C_*^2 \big) + \theta^2 \Big]^{\frac{1}{2}} \|\|u_{h,\tau}\|\|_{\mathcal{E}}, \tag{4.8}$$

with $\delta = C_S C_{\mathrm{inv}}$.

Proof. Let $u_{h,\tau} \in \mathcal{V}_h^\tau$. Using the triangle inequality and the definition of $v_{h,\tau}$, we get

$$\|\|v_{h,\tau}\|\|_{\mathcal{E}}^2 \leq 2\|\|\Pi_r^t(\varphi u_{h,\tau})\|\|_{\mathcal{E}}^2 + 2\theta^2 \|\|u_{h,\tau}\|\|_{\mathcal{E}}^2. \tag{4.9}$$

Hence, it only remains to bound the first term on the right-hand side of (4.9).

The following estimates follow immediately from the stability properties of Π_r^t, the commutativity of the spatial gradient ∇ and the operator Π_r^t, and bound (4.2a) for φ:

$$\|\Pi_r^t(\varphi u_{h,\tau})\|_{L^2(Q_T)}^2 \leq (eT)^2 \|u_{h,\tau}\|_{L^2(Q_T)}^2, \tag{4.10a}$$

$$\nu \|\nabla(\Pi_r^t(\varphi u_{h,\tau}))\|_{L^2(Q_T)^d}^2 \leq \nu(eT)^2 \|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2, \tag{4.10b}$$

$$|\varphi u_{h,\tau}|_{\mathrm{j}}^2 \leq (eT)^2 |u_{h,\tau}|_{\mathrm{j}}^2. \tag{4.10c}$$

By using the triangle inequality, estimates (4.4c) and (4.4d), bound (4.2a) for φ, and estimate (4.10c), we obtain

$$|\Pi_r^t(\varphi u_{h,\tau})|_{\mathrm{j}}^2 \leq 2|(\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})|_{\mathrm{j}}^2 + 2|\varphi u_{h,\tau}|_{\mathrm{j}}^2 \leq 2C_S^2 \tau \|u_{h,\tau}\|_{L^2(Q_T)}^2 + 2(eT)^2 |u_{h,\tau}|_{\mathrm{j}}^2. \tag{4.11}$$

We now bound the SUPG-seminorm of $\Pi_r^t(\varphi u_{h,\tau})$. For all $K \in \Omega_h$ and $n = 1, \ldots, N$, we have

$$|\Pi_r^t(\varphi u_{h,\tau})|_{K_n, \mathrm{supg}}^2 = \lambda_{K_n} \|\widetilde{\mathcal{L}}_{h,\tau}^K\big(\Pi_r^t(\varphi u_{h,\tau})\big)\|_{L^2(K_n)}^2 + \lambda_{K_n} \beta_{K_n}^2 \|\Pi_r^t\big(\varphi \nabla(\mathrm{Id} - \Pi_k^{\nabla,K}) u_{h,\tau}\big)\|_{L^2(K_n)^d}^2. \tag{4.12}$$

We consider the first term on the right-hand side of (4.12). Using the triangle inequality, the orthogonality properties of $\Pi_k^{0,K}$ and $\Pi_{k-1}^{0,K}$, estimates (4.4a) and (4.4b), the inverse estimate (4.3a), bounds (4.2a) and (4.2b), assumption (3.16) on the choice

of λ_{K_n}, and the mild condition (3.17), we get

$$
\lambda_{K_n} \|\widetilde{\mathcal{L}}_{h,\tau}^K \big(\Pi_r^t(\varphi u_{h,\tau})\big)\|_{L^2(K_n)}^2
$$

$$
\leq 2\lambda_{K_n}\|\widetilde{\mathcal{L}}_{h,\tau}^K(\varphi u_{h,\tau})\|_{L^2(K_n)}^2 + 2\lambda_{K_n}\|\widetilde{\mathcal{L}}_{h,\tau}^K(\mathrm{Id}-\Pi_r^t)(\varphi u_{h,\tau})\|_{L^2(K_n)}^2
$$

$$
\leq 2\lambda_{K_n}\|\partial_t(\varphi\Pi_k^{0,K}u_{h,\tau}) + \varphi\boldsymbol{\beta}\cdot\boldsymbol{\Pi}_{k-1}^{0,K}\nabla u_{h,\tau}\|_{L^2(K_n)}^2 + 4\lambda_{K_n}\|\partial_t(\mathrm{Id}-\Pi_r^t)(\varphi\Pi_k^{0,K}u_{h,\tau})\|_{L^2(K_n)}^2
$$

$$
+ 4\lambda_{K_n}\|\boldsymbol{\beta}\cdot(\mathrm{Id}-\Pi_r^t)(\varphi\boldsymbol{\Pi}_{k-1}^{0,K}\nabla u_{h,\tau})\|_{L^2(K_n)}^2
$$

$$
\leq 4\lambda_{K_n}\|\varphi'\Pi_k^{0,K}u_{h,\tau}\|_{L^2(K_n)}^2 + 4\lambda_{K_n}\|\varphi\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + 4\lambda_{K_n}\|\partial_t(\mathrm{Id}-\Pi_r^t)(\varphi\Pi_k^{0,K}u_{h,\tau})\|_{L^2(K_n)}^2
$$

$$
+ 4\lambda_{K_n}\|\boldsymbol{\beta}\cdot(\mathrm{Id}-\Pi_r^t)(\varphi\boldsymbol{\Pi}_{k-1}^{0,K}\nabla u_{h,\tau})\|_{L^2(K_n)}^2
$$

$$
\leq 4(1+C_S^2)\lambda_{K_n}\|u_{h,\tau}\|_{L^2(K_n)}^2 + 4(eT)^2\lambda_{K_n}\|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + 4\bar{\beta}_{Q_T}^2 C_S^2\lambda_{K_n}\tau_n^2\|\nabla u_{h,\tau}\|_{L^2(K_n)^d}^2
$$

$$
\leq 4\zeta(1+C_S^2)\bar{\beta}_{Q_T}^{-1}h_K\|u_{h,\tau}\|_{L^2(K_n)}^2 + 4(eT)^2\lambda_{K_n}\|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + 4\bar{\beta}_{Q_T}C_S^2 C_{\mathrm{inv}}^2\zeta\tau_n^2h_K^{-1}\|u_{h,\tau}\|_{L^2(K_n)}^2
$$

$$
\leq 4\zeta\big((1+C_S^2)\bar{\beta}_{Q_T}^{-1}h_K + \delta^2\bar{\beta}_{Q_T}C_*^2\big)\|u_{h,\tau}\|_{L^2(K_n)}^2 + 4(eT)^2\lambda_{K_n}\|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2, \tag{4.13}
$$

where $\delta = C_S C_{\mathrm{inv}}$.

Moreover, using the stability properties of Π_r^t and bound (4.2a), the second term on the right-hand side of (4.12) can be bounded as follows:

$$
\lambda_{K_n}\beta_{K_n}^2\|\Pi_r^t\big(\varphi\nabla(\mathrm{Id}-\Pi_k^{\nabla,K})u_{h,\tau}\big)\|_{L^2(K_n)^d}^2 \leq (eT)^2\lambda_{K_n}\beta_{K_n}^2\|\nabla(\mathrm{Id}-\Pi_k^{\nabla,K})u_{h,\tau}\|_{L^2(K_n)^d}^2,
$$

which, combined with (4.13) and summing up for all n, gives

$$
|\Pi_r^t(\varphi u_{h,\tau})|_{\mathrm{supg}}^2 \leq 4(eT)^2|u_{h,\tau}|_{\mathrm{supg}}^2 + 4\zeta\big((1+C_S^2)\bar{\beta}_{Q_T}^{-1}h + \delta^2\bar{\beta}_{Q_T}C_*^2\big)\|u_{h,\tau}\|_{L^2(Q_T)}^2. \tag{4.14}
$$

Therefore, combining bound (4.9) with estimates (4.10a), (4.10b), (4.11), and (4.14), we get the desired result. $\square$

Lemma 4.5. *Under the notation and assumptions of Lemma 4.4, the following bound holds:*

$$
\nu a_{h,\tau}(u_{h,\tau}, v_{h,\tau}) \geq \nu\alpha_*(T+\theta)\|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2,
$$

where α_* is the stability constant in (3.7).

Proof. Let $u_{h,\tau} \in \mathcal{V}_h^\tau$. By the stability property in (3.7) of the bilinear form $a_h(\cdot,\cdot)$, and bound (4.2a) for φ, we have

$$
\nu a_{h,\tau}(u_{h,\tau}, v_{h,\tau}) = \nu a_{h,\tau}(u_{h,\tau}, \varphi u_{h,\tau}) - \nu a_{h,\tau}(u_{h,\tau}, (\mathrm{Id}-\Pi_r^t)(\varphi u_{h,\tau})) + \theta\nu a_{h,\tau}(u_{h,\tau}, u_{h,\tau})
$$

$$
\geq (T+\theta)\alpha_*\nu\|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2 - \nu a_{h,\tau}(u_{h,\tau}, (\mathrm{Id}-\Pi_r^t)(\varphi u_{h,\tau})). \tag{4.15}
$$

The last term on the right-hand side of (4.15) vanishes due to the orthogonality properties of Π_r^t and Lemma 4.3, which completes the proof. $\square$

Lemma 4.6. *Under the notation and assumptions of Lemma 4.4, the following bound holds:*

$$
b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, v_{h,\tau}) \geq -\delta_t C_*^2\|u_{h,\tau}\|_{L^2(Q_T)}^2, \tag{4.16}
$$

with $\delta_t = \delta C_0\|\partial_t\boldsymbol{\beta}\|_{L^\infty(Q_T)^d}$, and C_* as in the mild condition (3.17).

Proof. Let $u_{h,\tau} \in \mathcal{V}_h^\tau$. Using the skew symmetry of the bilinear form $b_{h,\tau}^{\mathrm{skew}}(\cdot,\cdot)$, and the commutativity of $\boldsymbol{\Pi}_k^{0,K}$ and Π_r^t, we have

$$
b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, v_{h,\tau}) = b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, \varphi u_{h,\tau}) - b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, (\mathrm{Id}-\Pi_r^t)(\varphi u_{h,\tau})) + \theta b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, u_{h,\tau})
$$

$$
= -b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, (\mathrm{Id}-\Pi_r^t)(\varphi u_{h,\tau}))
$$

$$
= -\frac{1}{2}\sum_{n=1}^N\sum_{K\in\Omega_h}\Big[\big(\boldsymbol{\beta}\cdot\boldsymbol{\Pi}_k^{0,K}\nabla u_{h,\tau},\ (\mathrm{Id}-\Pi_r^t)(\varphi\Pi_k^{0,K}u_{h,\tau})\big)_{K_n} - \big((\mathrm{Id}-\Pi_r^t)(\varphi\boldsymbol{\Pi}_k^{0,K}\nabla u_{h,\tau}),\ \boldsymbol{\beta}\Pi_k^{0,K}u_{h,\tau}\big)_{K_n}\Big]. \tag{4.17}
$$

In particular, if $\boldsymbol{\beta} = \boldsymbol{\beta}(\boldsymbol{x})$, then $b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, v_{h,\tau}) = 0$. Otherwise, since $\Pi_k^t(\boldsymbol{\beta})\cdot\boldsymbol{\Pi}_k^{0,K}\nabla u_{h,\tau} \in \mathbb{P}_r(\mathcal{T}_\tau; L^2(\Omega))$ and $\Pi_0^t(\boldsymbol{\beta})\Pi_k^{0,K}u_{h,\tau} \in \mathbb{P}_r(\mathcal{T}_\tau; L^2(\Omega)^d)$, using identity (4.17) and the orthogonality properties of Π_r^t, we obtain

$$
b_{h,\tau}^{\mathrm{skew}}(u_{h,\tau}, v_{h,\tau}) = -\frac{1}{2}\sum_{n=1}^N\sum_{K\in\Omega_h}\Big[\big((\mathrm{Id}-\Pi_0^t)\boldsymbol{\beta}\cdot\boldsymbol{\Pi}_k^{0,K}\nabla u_{h,\tau},\ (\mathrm{Id}-\Pi_r^t)(\varphi\Pi_k^{0,K}u_{h,\tau})\big)_{K_n}
$$

$$
- \big((\mathrm{Id}-\Pi_r^t)(\varphi\boldsymbol{\Pi}_k^{0,K}\nabla u_{h,\tau}),\ (\mathrm{Id}-\Pi_0^t)\boldsymbol{\beta}\Pi_k^{0,K}u_{h,\tau}\big)_{K_n}\Big]. \tag{4.18}
$$

For any $K \in \Omega_h$ and $n = 1, \ldots, N$, using the stability properties of $\Pi_k^{0,K}$, the local inverse estimate in (4.3a), estimate (4.4a), error bound (4.6) for β, and the Hölder inequality, we get

$$
\begin{aligned}
-\big((\mathrm{Id} - \Pi_0^t)\beta \cdot \Pi_k^{0,K}\nabla u_{h,\tau}, \ (\mathrm{Id} - \Pi_r^t)(\varphi\Pi_k^{0,K}u_{h,\tau})\big)_{K_n} \\
\geq -\|\beta - \Pi_0^t\beta\|_{L^\infty(K_n)^d} \|(\mathrm{Id} - \Pi_r^t)(\varphi\Pi_k^{0,K}u_{h,\tau})\Pi_k^{0,K}\nabla u_{h,\tau}\|_{L^1(K_n)^d} \\
\geq -\|\beta - \Pi_0^t\beta\|_{L^\infty(K_n)^d} \|(\mathrm{Id} - \Pi_r^t)(\varphi\Pi_k^{0,K}u_{h,\tau})\|_{L^2(K_n)} \|\Pi_k^{0,K}\nabla u_{h,\tau}\|_{L^2(K_n)^d} \\
\geq -\delta_t h_K^{-1}\tau_n^2 \|u_{h,\tau}\|_{L^2(K_n)}^2,
\end{aligned}
\tag{4.19}
$$

where $\delta_t = \delta C_0 \|\partial_t\beta\|_{L^\infty(Q_T)^d}$.

The following bound can be proven in a similar way:

$$
-\big((\mathrm{Id} - \Pi_r^t)(\varphi\Pi_k^{0,K}\nabla u_{h,\tau}), \ (\mathrm{Id} - \Pi_0^t)\beta\Pi_k^{0,K}u_{h,\tau}\big)_{K_n} \geq -\delta_t h_K^{-1}\tau_n^2 \|u_{h,\tau}\|_{L^2(K_n)}^2,
$$

which, combined with identity (4.18), the mild condition (3.17) and bound (4.19), leads to (4.16). $\square$

Lemma 4.7. *Under the assumptions of Lemma 4.4, the following bound holds:*

$$
m_{h,\tau}^{\partial_t}(u_{h,\tau}, v_{h,\tau}) \geq \mu_* \left(\left(\frac{1}{2} - \left(\frac{\mu^*}{\mu_*}\right)^2 \frac{C_S^2\tau}{2\theta} \right) \|u_{h,\tau}\|_{L^2(Q_T)}^2 + T|u_{h,\tau}|_{\mathrm{j}}^2 \right),
\tag{4.20}
$$

where μ_* and μ^* are the stability constants in (3.4), and C_S is the constant from Lemma 4.2.

Proof. Let $u_{h,\tau} \in \mathcal{V}_h^\tau$. By adding and subtracting suitable terms, we have

$$
\begin{aligned}
m_{h,\tau}^{\partial_t}(u_{h,\tau}, v_{h,\tau}) &= m_{h,\tau}^{\partial_t}(u_{h,\tau}, \varphi u_{h,\tau}) - m_{h,\tau}^{\partial_t}(u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})) + \theta m_{h,\tau}^{\partial_t}(u_{h,\tau}, u_{h,\tau}) \\
&=: M_1 + M_2 + M_3.
\end{aligned}
\tag{4.21}
$$

We treat each term M_i, $i = 1, 2, 3$, separately.

Bound for M_1. Since $\varphi'(t) = -(1/T)\varphi(t)$ for all $t \in [0, T]$, the following identity follows:

$$
\varphi\partial_t(u_{h,\tau}^2) = (1/T)u_{h,\tau}^2\varphi + \partial_t(\varphi u_{h,\tau}^2),
$$

which, together with the identity $\frac{1}{2}[\![\varphi w^2]\!]_n - [\![\varphi w]\!]_n w(\cdot, t_n^+) = \frac{1}{2}[\![\sqrt{\varphi}w]\!]_n^2$, the symmetry and stability of the bilinear form $m_h(\cdot, \cdot)$, and bound (4.2a) for φ, leads to

$$
\begin{aligned}
M_1 &= m_{h,\tau}^{\partial_t}(u_{h,\tau}, \varphi u_{h,\tau}) \\
&= \frac{1}{2}\sum_{n=1}^N \int_{I_n} \varphi(t)\frac{\mathrm{d}}{\mathrm{d}t}m_h(u_{h,\tau}, u_{h,\tau})\,\mathrm{d}t - \sum_{n=1}^{N-1} m_h([\![u_{h,\tau}]\!]_n, \varphi(t_n)u_{h,\tau}(\cdot, t_n^+)) + m_h(u_{h,\tau}(\cdot, 0), \varphi(0)u_{h,\tau}(\cdot, 0)) \\
&= \frac{1}{2T}\sum_{n=1}^N \int_{I_n} \varphi(t)m_h(u_{h,\tau}, u_{h,\tau})\,\mathrm{d}t + \frac{1}{2}\sum_{n=1}^N \int_{I_n} \frac{\mathrm{d}}{\mathrm{d}t}\Big(\varphi(t)m_h(u_{h,\tau}, u_{h,\tau})\Big)\,\mathrm{d}t \\
&\quad - \sum_{n=1}^{N-1} m_h([\![u_{h,\tau}]\!]_n, \varphi(t_n)u_{h,\tau}(\cdot, t_n^+)) + m_h(u_{h,\tau}(\cdot, 0), \varphi(0)u_{h,\tau}(\cdot, 0)) \\
&\geq \frac{\mu_*}{2T}\|\sqrt{\varphi}u_{h,\tau}\|_{L^2(Q_T)}^2 + \frac{1}{2}\Big[m_h(\varphi(T)u_{h,\tau}(\cdot, T), u_{h,\tau}(\cdot, T)) + \sum_{n=1}^{N-1}[\![\varphi(t_n)m_h(u_{h,\tau}, u_{h,\tau})]\!]_n - m_h(\varphi(0)u_{h,\tau}(\cdot, 0), u_{h,\tau}(\cdot, 0)) \Big] \\
&\quad - \sum_{n=1}^{N-1} m_h([\![u_{h,\tau}]\!]_n, \varphi(t_n)u_{h,\tau}(\cdot, t_n^+)) + m_h(\varphi(0)u_{h,\tau}(\cdot, 0), u_{h,\tau}(\cdot, 0)) \\
&= \frac{\mu_*}{2T}\|\sqrt{\varphi}u_{h,\tau}\|_{L^2(Q_T)}^2 + \frac{1}{2}\Big[m_h(\varphi(T)u_{h,\tau}(\cdot, T), u_{h,\tau}(\cdot, T)) + \sum_{n=1}^{N-1} m_h(\varphi(t_n)[\![u_{h,\tau}]\!]_n, [\![u_{h,\tau}]\!]_n) + m_h(\varphi(0)u_{h,\tau}(\cdot, 0), u_{h,\tau}(\cdot, 0)) \Big] \\
&\geq \mu_* \left(\frac{1}{2T}\|\sqrt{\varphi}u_{h,\tau}\|_{L^2(Q_T)}^2 + |\sqrt{\varphi}u_{h,\tau}|_{\mathrm{j}}^2 \right) \\
&\geq \mu_* \left(\frac{1}{2}\|u_{h,\tau}\|_{L^2(Q_T)}^2 + T|u_{h,\tau}|_{\mathrm{j}}^2 \right).
\end{aligned}
\tag{4.22}
$$

Bound for M_2. Using the orthogonality properties of Π_r^t, Lemma 4.3, estimate (4.4c), the local inverse estimate in (4.3b), the symmetry and stability of the bilinear form $m_h(\cdot, \cdot)$, and the Young inequality, we obtain

$$
\begin{aligned}
M_2 &= -m_{h,\tau}^{\partial_t}(u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})) \\
&= -\sum_{n=1}^N \sum_{K \in \Omega_h} \Big[(\partial_t u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau}))_{K_n} + \int_{I_n} s_m^K\big((\mathrm{Id} - \Pi_k^{0,K})\partial_t u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi(\mathrm{Id} - \Pi_k^{0,K})u_{h,\tau})\big)\,\mathrm{d}t \Big]
\end{aligned}
$$

$$+ \sum_{n=1}^{N-1} m_h\big(\llbracket u_{h,\tau} \rrbracket_n, (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})(\cdot, t_n^+)\big) - m_h\big(u_{h,\tau}(\cdot, 0), (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})(\cdot, 0)\big)$$

$$= \sum_{n=1}^{N-1} m_h\big(\llbracket u_{h,\tau} \rrbracket_n, (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})(\cdot, t_n^+)\big) - m_h\big(u_{h,\tau}(\cdot, 0), (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})(\cdot, 0)\big)$$

$$\geq -\mu^* \left(\sum_{n=1}^{N-1} \|\llbracket u_{h,\tau} \rrbracket_n\|_{L^2(\Omega)} \|(\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})\|_{L^2(\Sigma_n)} + \|u_{h,\tau}\|_{L^2(\Sigma_0)} \|(\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})\|_{L^2(\Sigma_0)} \right)$$

$$\geq -\mu^* \left(\theta \frac{\mu_*}{\mu^*} |u_{h,\tau}|_{\mathrm{j}}^2 + \frac{C_S^2 \tau}{2\theta} \cdot \frac{\mu^*}{\mu_*} \|u_{h,\tau}\|_{L^2(Q^T)}^2 \right). \tag{4.23}$$

Bound for M_3. Integration by parts in time, the symmetry and stability of the bilinear form $m_{h,\tau}^{\partial_t}(\cdot, \cdot)$, and the identity $\frac{1}{2}\llbracket w^2 \rrbracket_n - \llbracket w \rrbracket_n w(\cdot, t_n^+) = \frac{1}{2}\llbracket w \rrbracket_n^2$ yield

$$M_3 = \theta m_{h,\tau}^{\partial_t}(u_{h,\tau}, u_{h,\tau})$$

$$= \frac{\theta}{2} \sum_{n=1}^{N} \int_{I_n} \frac{\mathrm{d}}{\mathrm{d}t} m_h(u_{h,\tau}, u_{h,\tau}) \, \mathrm{d}t - \theta \sum_{n=1}^{N-1} m_h\big(\llbracket u_{h,\tau} \rrbracket_n, u_{h,\tau}(\cdot, t_n^+)\big) + \theta m_h\big(u_{h,\tau}(\cdot, 0), u_{h,\tau}(\cdot, 0)\big)$$

$$= \frac{\theta}{2} \left(m_h\big(u_{h,\tau}(\cdot, T), u_{h,\tau}(\cdot, T)\big) + \sum_{n=1}^{N-1} \llbracket m_h(u_{h,\tau}, u_{h,\tau}) \rrbracket_n - m_h\big(u_{h,\tau}(\cdot, 0), u_{h,\tau}(\cdot, 0)\big) \right)$$

$$\quad - \theta \sum_{n=1}^{N-1} m_h\big(\llbracket u_{h,\tau} \rrbracket_n, u_{h,\tau}(\cdot, t_n^+)\big) + \theta m_h\big(u_{h,\tau}(\cdot, 0), u_{h,\tau}(\cdot, 0)\big)$$

$$= \frac{\theta}{2} \left(m_h\big(u_{h,\tau}(\cdot, T), u_{h,\tau}(\cdot, T)\big) + \sum_{n=1}^{N-1} m_h\big(\llbracket u_{h,\tau} \rrbracket_n, \llbracket u_{h,\tau} \rrbracket_n\big) + m_h\big(u_{h,\tau}(\cdot, 0), u_{h,\tau}(\cdot, 0)\big) \right)$$

$$\geq \theta \mu_* |u_{h,\tau}|_{\mathrm{j}}^2. \tag{4.24}$$

Combining identity (4.21) with bounds (4.22), (4.23), and (4.24), we get (4.20). $\square$

Lemma 4.8. *Under the assumptions of Lemma 4.4, the following bound holds:*

$$s_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, v_{h,\tau}) \geq \frac{\alpha_* T}{2} |u_{h,\tau}|_{\mathrm{supg}}^2 - \zeta \left[\frac{h}{\bar{\beta}_{Q_T}} \Big(\frac{1}{T} + \frac{1}{2\theta}\Big) + \frac{1}{\alpha_* \theta} \big(C_S^2 \bar{\beta}_{Q_T}^{-1} h + \delta^2 \bar{\beta}_{Q_T} C_*^2\big) \right] \|u_{h,\tau}\|_{L^2(Q_T)}^2$$

$$\quad - \zeta \left[e^2 T + \Big(\frac{\alpha_*}{2} + 1\Big)\theta \right] \nu \|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2, \tag{4.25}$$

with $\delta = C_S C_{\mathrm{inv}}$, α_* is the stability constant in (3.6), and C_* is the constant in the mild condition (3.17).

Proof. Let $u_{h,\tau} \in \mathcal{V}_h^\tau$. By adding and subtracting suitable terms, we have

$$s_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, v_{h,\tau}) = s_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, \varphi u_{h,\tau}) - s_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau})) + \theta s_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, u_{h,\tau})$$

$$=: J_1 + J_2 + J_3. \tag{4.26}$$

We bound each term J_i, $i = 1, 2, 3$, separately.

Bound for J_1. Using the triangle and the Young inequalities, the stability property in (3.6) of $s_a^K(\cdot, \cdot)$, the inverse estimates in (4.3b) and (4.3c), bounds (4.2a) and (4.2b) for φ, the stability properties of $\Pi_k^{0,K}$ and $\Pi_{k-1}^{0,K}$, and assumption (3.16) on the choice of λ_{K_n}, we get

$$J_1 = s_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, \varphi u_{h,\tau})$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \lambda_{K_n} \left[\Big(\mathcal{L}_{h,\tau}^K u_{h,\tau}, \widetilde{\mathcal{L}}_{h,\tau}^K(\varphi u_{h,\tau})\Big)_{K_n} + \beta_{K_n}^2 \int_{I_n} \varphi s_a^K\big((\mathrm{Id} - \Pi_k^{\nabla,K})u_{h,\tau}, (\mathrm{Id} - \Pi_k^{\nabla,K})u_{h,\tau}\big) \, \mathrm{d}t \right]$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\lambda_{K_n}\Big(\mathcal{L}_{h,\tau}^K u_{h,\tau}, \partial_t(\varphi \Pi_k^{0,K} u_{h,\tau}) + \varphi \boldsymbol{\beta} \cdot \Pi_{k-1}^{0,K} \nabla u_{h,\tau}\Big)_{K_n} + \beta_{K_n}^2 \lambda_{K_n} \int_{I_n} \varphi s_a^K\big((\mathrm{Id} - \Pi_k^{\nabla,K})u_{h,\tau}, (\mathrm{Id} - \Pi_k^{\nabla,K})u_{h,\tau}\big) \, \mathrm{d}t \right]$$

$$\geq \sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\lambda_{K_n}\Big(\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}, \varphi' \Pi_k^{0,K} u_{h,\tau}\Big)_{K_n} + \lambda_{K_n} \|\sqrt{\varphi}\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 \right.$$

$$\left. + \check{\alpha} \beta_{K_n}^2 \lambda_{K_n} \|\sqrt{\varphi}\nabla(\mathrm{Id} - \Pi_k^{\nabla,K})u_{h,\tau}\|_{L^2(K_n)^d}^2 - \nu \lambda_{K_n}\Big(\mathrm{div}\, \boldsymbol{\Pi}_{k-1}^{0,K} \nabla u_{h,\tau}, \varphi' \Pi_k^{0,K} u_{h,\tau} + \varphi \widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\Big)_{K_n} \right]$$

$$\geq \sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\frac{T}{2} \lambda_{K_n} \|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + \breve{a} T \lambda_{K_n} \beta_{K_n}^2 \|\nabla(\mathrm{Id} - \Pi_k^{\nabla,K}) u_{h,\tau}\|_{L^2(K_n)^d}^2 \right.$$

$$\left. - \lambda_{K_n} \left(\frac{1}{T} + \frac{1}{2\theta} \right) \|\varphi' \Pi_k^{0,K} u_{h,\tau}\|_{L^2(K_n)}^2 - \nu^2 \lambda_{K_n} \left(\frac{\theta}{2} + e^2 T \right) \|\operatorname{div} \boldsymbol{\Pi}_{k-1}^{0,K} \nabla u_{h,\tau}\|_{L^2(K_n)}^2 \right]$$

$$\geq \sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\frac{T}{2} \lambda_{K_n} \|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + (\breve{a} T) \lambda_{K_n} \beta_{K_n}^2 \|\nabla(\mathrm{Id} - \Pi_k^{\nabla,K}) u_{h,\tau}\|_{L^2(K_n)^d}^2 \right.$$

$$\left. - \lambda_{K_n} \left(\frac{1}{T} + \frac{1}{2\theta} \right) \|\varphi' \Pi_k^{0,K} u_{h,\tau}\|_{L^2(K_n)}^2 - \nu^2 \lambda_{K_n} \left(\frac{\theta}{2} + e^2 T \right) C_{\mathrm{inv}}^2 h_K^{-2} \|\boldsymbol{\Pi}_{k-1}^{0,K} \nabla u_{h,\tau}\|_{L^2(K_n)}^2 \right]$$

$$\geq \frac{T \alpha_*}{2} |u_{h,\tau}|_{\mathrm{supg}}^2 - \frac{\zeta h}{\bar{\beta}_{Q_T}} \left(\frac{1}{T} + \frac{1}{2\theta} \right) \|u_{h,\tau}\|_{L^2(Q_T)}^2 - \nu \zeta \left(\frac{\theta}{2} + e^2 T \right) \|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2. \tag{4.27}$$

Bound for J_2. The following identity follows from Lemma 4.3

$$J_2 = -s_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau}))$$

$$= -\sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\lambda_{K_n} \left(\mathcal{L}_{h,\tau}^K u_{h,\tau}, \widetilde{\mathcal{L}}_{h,\tau}^K (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau}) \right)_{K_n} + \beta_{K_n}^2 \lambda_{K_n} \int_{I_n} s_a^K \left((\mathrm{Id} - \Pi_k^{\nabla,K}) u_{h,\tau}, (\mathrm{Id} - \Pi_r^t)(\varphi(\mathrm{Id} - \Pi_k^{\nabla,K}) u_{h,\tau}) \right) \mathrm{d}t \right]$$

$$= -\sum_{n=1}^{N} \sum_{K \in \Omega_h} \lambda_{K_n} \left(\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau} - \nu \operatorname{div} \boldsymbol{\Pi}_{k-1}^{0,K} \nabla u_{h,\tau}, \widetilde{\mathcal{L}}_{h,\tau}^K (\mathrm{Id} - \Pi_r^t)(\varphi u_{h,\tau}) \right)_{K_n}.$$

Using the Cauchy–Schwarz, the triangle, and the Young inequalities, estimates (4.4a) and (4.4b), the local inverse estimates in Lemma 4.1, assumption (3.16) on the choice of λ_{K_n}, and condition (3.17), we get

$$J_2 \geq -\sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\frac{\lambda_{K_n} \alpha_* \theta}{2} \|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + \frac{\nu^2 \lambda_{K_n} \alpha_* \theta}{2} \|\operatorname{div} \boldsymbol{\Pi}_{k-1}^{0,K} \nabla u_{h,\tau}\|_{L^2(K_n)}^2 \right.$$

$$\left. + \frac{2\lambda_{K_n}}{\theta \alpha_*} \left(\|\partial_t(\mathrm{Id} - \Pi_r^t) \Pi_k^{0,K}(\varphi u_{h,\tau})\|_{L^2(K_n)}^2 + \|\boldsymbol{\beta} \cdot (\mathrm{Id} - \Pi_r^t)(\varphi \boldsymbol{\Pi}_{k-1}^{0,K} \nabla u_{h,\tau})\|_{L^2(K_n)}^2 \right) \right]$$

$$\geq -\sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\frac{\lambda_{K_n} \alpha_* \theta}{2} \|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + \frac{\nu^2 \lambda_{K_n} \alpha_* \theta}{2} C_{\mathrm{inv}}^2 h_K^{-2} \|\nabla u_{h,\tau}\|_{L^2(K_n)^d}^2 \right.$$

$$\left. + \frac{2\lambda_{K_n}}{\theta \alpha_*} \left(C_S^2 + \delta^2 \bar{\beta}_{Q_T}^2 \tau_n^2 h_K^{-2} \right) \|u_{h,\tau}\|_{L^2(K_n)}^2 \right]$$

$$\geq -\sum_{n=1}^{N} \sum_{K \in \Omega_h} \left[\frac{\lambda_{K_n} \alpha_* \theta}{2} \|\widetilde{\mathcal{L}}_{h,\tau}^K u_{h,\tau}\|_{L^2(K_n)}^2 + \frac{\nu \zeta \alpha_* \theta}{2} \|\nabla u_{h,\tau}\|_{L^2(K_n)^d}^2 \right.$$

$$\left. + \frac{2\zeta}{\theta \alpha_*} \left(C_S^2 \bar{\beta}_{Q_T}^{-1} h_K + \delta^2 \bar{\beta}_{Q_T} \tau_n^2 h_K^{-1} \right) \|u_{h,\tau}\|_{L^2(K_n)}^2 \right]$$

$$\geq -\frac{\alpha_* \theta}{2} |u_{h,\tau}|_{\mathrm{supg}}^2 - \frac{\nu \zeta \alpha_* \theta}{2} \|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2 - \frac{2\zeta}{\alpha_* \theta} \left(C_S^2 \bar{\beta}_{Q_T}^{-1} h + \delta^2 \bar{\beta}_{Q_T} C_*^2 \right) \|u_{h,\tau}\|_{L^2(Q_T)}^2, \tag{4.28}$$

with $\delta = C_S C_{\mathrm{inv}}$.

Bound for J_3. Similarly as for the bound for J_1, it can be shown that

$$J_3 \geq \frac{\alpha_* \theta}{2} |u_{h,\tau}|_{\mathrm{supg}}^2 - \frac{\nu \zeta \theta}{2} \|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2. \tag{4.29}$$

Combining identity (4.26) with bounds (4.27), (4.28), and (4.29), we get (4.25) □

We are now in a position to prove the stability and well-posedness of the method.

Proof of Theorem 3.6. Let $u_{h,\tau} \in \mathcal{V}_h^\tau$ and $v_{h,\tau} := \Pi_r^t(\varphi u_{h,\tau}) + \theta u_{h,\tau} \in \mathcal{V}_h^\tau$, with φ defined in (4.1) and the real $\theta > 0$ to be fixed later. From Lemmas 4.5, 4.6, 4.7, and 4.8, we get

$$\mathcal{B}_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, v_{h,\tau}) \geq \left[\mu_* \left(\frac{1}{2} - \left(\frac{\mu^*}{\mu_*} \right)^2 \frac{C_S^2 \tau}{2\theta} \right) - \delta_t C_*^2 - \zeta \left(\frac{h}{\bar{\beta}_{Q_T}} \left(\frac{1}{T} + \frac{1}{2\theta} \right) + \frac{2}{\theta \alpha_*} \left(C_S^2 \bar{\beta}_{Q_T}^{-1} h + \delta^2 \bar{\beta}_{Q_T} C_*^2 \right) \right) \right] \|u_{h,\tau}\|_{L^2(Q_T)}^2$$

$$+ \mu_* T |u_{h,\tau}|_j^2 + \frac{\alpha_* T}{2} |u_{h,\tau}|_{\mathrm{supg}}^2 + \left[T \left(\alpha_* - \zeta e^2 \right) + \theta \left(\alpha_* - \zeta \left(\frac{\alpha_*}{2} + 1 \right) \right) \right] \nu \|\nabla u_{h,\tau}\|_{L^2(Q_T)^d}^2.$$

Choosing ζ and C_* small enough, and θ large enough, we deduce that there exists a constant $\eta^* > 0$ independent of h, τ, and ν such that

$$\mathcal{B}_{h,\tau}^{\mathrm{supg}}(u_{h,\tau}, v_{h,\tau}) \geq \eta^* \|\|u_{h,\tau}\|\|_\varepsilon^2. \tag{4.30}$$

which, combined with Lemma 4.4, completes the proof.

Proof of Corollary 3.7. The result follows from the inf–sup stability estimate in Theorem 3.6 and the following continuity bound for the linear functional $\ell_{h,\tau}^{\text{supg}}(\cdot)$ in (3.12): for all $v_{h,\tau} \in \mathcal{V}_h^\tau$, it holds

$$|\ell_{h,\tau}^{\text{supg}}(v_{h,\tau})| \leq \sqrt{2}\left(\|u_0\|_{L^2(\Omega)} + \|f\|_{L^2(Q_T)}\right)\left(\|v_{h,\tau}\|_{L^2(Q_T)}^2 + \frac{1}{2}\|v_{h,\tau}\|_{L^2(\Sigma_0)}^2 + \sum_{n=1}^{N}\sum_{K \in \Omega_h}\lambda_{K_n}^2\|\widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau}\|_{L^2(K_n)}^2\right)^{\frac{1}{2}}$$

$$\leq \sqrt{2}(\|u_0\|_{L^2(\Omega)} + \|f\|_{L^2(Q_T)})|\!|\!|v_{h,\tau}|\!|\!|_{\mathcal{E}}.$$

Remark 4.9 (*Inf–Sup Constant for Long-Time Simulations*). The inf–sup stability constant γ_I is of the form

$$\gamma_I = \frac{\eta_*}{\eta^*},$$

where η^* is the constant in bound (4.30) and η_* is an upper bound for the constant on the right-hand side of (4.8).
Assuming that $T \gg 1$ and

$$\tau < \frac{1}{C_S^2}\left(\frac{\mu^*}{\mu_*}\right)^2, \quad h < \bar{\beta}_{Q_T}\min\left\{\frac{\alpha_*}{2C_S^2},\, 1\right\}, \quad C_*^2 < \min\left\{\frac{\mu_*}{8\delta_t},\, \frac{\alpha_*}{2\delta^2\bar{\beta}_{Q_T}}\right\}, \quad \zeta < \min\left\{\frac{\alpha_*}{2e^2},\, \frac{\alpha_*}{\alpha_*+2},\, \frac{\mu^*T}{32}\right\}, \quad \theta = \max\left\{4,\, \frac{T}{2}\right\},$$

we have that

$$\eta^* \geq \min\left\{\frac{\mu_*}{8},\, \mu_*T,\, \alpha_*T\right\} = \frac{\mu_*}{8} \quad \text{and} \quad \eta_* = \mathcal{O}(T).$$

Therefore, the inf–sup stability constant $\gamma_I \sim T$.   ∎

5. Convergence analysis

In the present section, we prove Theorem 3.10, which will follow as a simplified case of the more general error bounds derived below. We start by recalling some polynomial and VE approximation results. The Bramble–Hilbert lemma (see e.g., [35, Lemma 4.3.8]) implies the following approximation properties of the polynomial projections introduced in Section 3.2.

Lemma 5.1 (*Estimates for Polynomial Projections*). *Under Assumption (A1), for any element $K \in \Omega_h$ and any sufficiently smooth function ϕ defined on K, the following estimates hold:*

$$\|(\text{Id} - \Pi_k^{0,K})\phi\|_{W^{m,p}(K)} \lesssim h_K^{s-m}|\phi|_{W^{s,p}(K)} \qquad\qquad s, m \in \mathbb{N},\ m \leq s \leq k+1,\ p \in [1, +\infty], \tag{5.1a}$$

$$\|(\text{Id} - \Pi_k^{\nabla,K})\phi\|_{H^m(K)} \lesssim h_K^{s-m}|\phi|_{H^s(K)} \qquad\qquad s, m \in \mathbb{N},\ m \leq s \leq k+1,\ s \geq 1. \tag{5.1b}$$

Moreover, for any $I_n \in \mathcal{T}_\tau$ and for any sufficiently smooth function ψ defined on I_n, the $L^2(\mathcal{T}_\tau)$-orthogonal projection in $\mathbb{P}_r(\mathcal{T}_\tau)$ of ψ satisfies

$$\|(\text{Id} - \Pi_r^t)\psi\|_{H^m(I_n)} \lesssim \tau_n^{\ell-m}|\psi|_{H^\ell(I_n)} \qquad \ell, m \in \mathbb{N},\ m \leq \ell \leq r+1. \tag{5.2}$$

The next lemma concerns the optimal approximation properties of the VE space $\mathcal{V}_h$, see [26, Lemma 3.15 (for 2D) and §5.2 (for 3D)] and [38, Thm. 11] for more details; see instead [32] for the case of serendipity VE spaces.

Lemma 5.2 (*Approximation by VE Functions*). *Under Assumption (A1), for any $v \in H_0^1(\Omega) \cap H^{s+1}(\Omega_h)$ ($0 < s \leq k+1$), there exists $v_I \in \mathcal{V}_h$ such that, for all $K \in \Omega_h$, it holds*

$$\|v - v_I\|_{L^2(K)} + h_K\|\nabla(v - v_I)\|_{L^2(K)^d} \lesssim h_K^{s+1}|v|_{H^{s+1}(K)}. \tag{5.3}$$

The last ingredients are a standard trace inequality in one dimension, a stability bound for Π_r^t in the $H^1(I_n)$-seminorm, and a scaled Poincaré–Friedrichs inequality on polytopes.

Lemma 5.3 (*Trace Inequality*). *Let $I_n \in \mathcal{T}_\tau$. For any $\psi \in H^1(I_n)$, it holds*

$$|\psi(t_{n-1})|^2 + |\psi(t_n)|^2 \lesssim \tau_n^{-1}\|\psi\|_{L^2(I_n)}^2 + \tau_n\|\partial_t\psi\|_{L^2(I_n)}^2. \tag{5.4}$$

Lemma 5.4 (*Stability of Π_r^t*). *Let $I_n \in \mathcal{T}_\tau$. For any $\psi \in H^1(I_n)$, the following bound holds:*

$$\|\partial_t\Pi_r^t\psi\|_{L^2(I_n)} \lesssim \|\partial_t\psi\|_{L^2(I_n)}.$$

Proof. Let $I_n \in \mathcal{T}_\tau$ and $\psi \in H^1(I_n)$. Using the polynomial inverse estimate (4.3b), the stability properties of Π_r^t, and the standard Poincaré inequality, we get

$$\|\partial_t \Pi_r^t \psi\|_{L^2(I_n)} = \|\partial_t (\Pi_r^t \psi - \Pi_0^t \psi)\|_{L^2(I_n)} \lesssim \tau_n^{-1} \|\Pi_r^t (\psi - \Pi_0^t \psi)\|_{L^2(I_n)} \lesssim \tau_n^{-1} \|(\mathrm{Id} - \Pi_0^t)\psi\|_{L^2(I_n)} \lesssim \|\partial_t \psi\|_{L^2(I_n)},$$

which completes the proof. $\square$

Lemma 5.5 (*Scaled Poincaré–Friedrichs Inequality (See e.g., [39, Lemma 2.2])*). *Under Assumption 3.1, for any element $K \in \Omega_h$ and $v \in H^1(K)$, it holds*

$$\|(\mathrm{Id} - \Pi_k^{\nabla,K})v\|_{L^2(K)} \lesssim h_K \|\nabla(\mathrm{Id} - \Pi_k^{\nabla,K})v\|_{L^2(K)^d}.$$

In the forthcoming convergence analysis, estimates (5.1a), (5.1b), and (5.3), as well as the scaled Poincaré–Friedrichs inequality in Lemma 5.5, are applied pointwise in time, whereas estimate (5.2) and the trace inequality (5.4) are applied pointwise in space.

5.1. Some preliminary assumptions and notations

Henceforth, we assume that the solution u to the continuous weak formulation (2.2) has the following parabolic regularity:

$$u \in L^2(0,T;H^2(\Omega)) \cap L^\infty(0,T;H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega)). \tag{5.5}$$

Consequently, a density argument can be used to show that

$$(\partial_t u - \nu \Delta u + \boldsymbol{\beta} \cdot \nabla u, v)_{Q_T} = (f, v)_{Q_T} \qquad \forall v \in L^2(Q_T). \tag{5.6}$$

For convenience, we also define the following operators:

$$\mathcal{L}^K u := \partial_t u - \nu \Delta u + \boldsymbol{\beta} \cdot \nabla u, \qquad \widetilde{\mathcal{L}}^K v := \partial_t v + \boldsymbol{\beta} \cdot \nabla v, \tag{5.7}$$

and the following bilinear forms and linear functionals:

$$m^{\partial_t}(u,v) := (\partial_t u, v)_{Q_T} + (u, v)_{\Sigma_0}, \qquad a^T(u,v) := \int_0^T a(u,v)\,\mathrm{d}t,$$

$$b^T(u,v) := \int_0^T b(u,v)\,\mathrm{d}t, \qquad s^{\mathrm{supg}}(u,v) := \sum_{K \in \Omega_h} \sum_{n=1}^N \lambda_{K_n} (\mathcal{L}^K u, \widetilde{\mathcal{L}}^K v)_{K_n},$$

$$\ell_f(v) := (f, v)_{Q_T}, \qquad \ell_f^{\mathrm{supg}}(v) := \sum_{K \in \Omega_h} \sum_{n=1}^N (f, \lambda_{K_n} \widetilde{\mathcal{L}}^K v)_{K_n},$$

$$\ell_{u_0}(v) := (u_0, v)_\Omega.$$

Remark 5.6 (*Parabolic Regularity*). The parabolic regularity assumption (5.5) holds, for instance, if Ω is convex, compatibility of the initial and boundary conditions is satisfied (i.e., $u_0 \in H_0^1(\Omega)$), and $\boldsymbol{\beta}$ does not depend on t (see [24, Thm. 5 in Ch. 7.1] for smooth domains, which can be extended to convex domains using the results in [40, Ch. 3]) $\blacksquare$

5.2. A priori error estimates

Let the solution u to the continuous weak formulation (2.2) satisfy the parabolic regularity (5.5), and let $u_{h,\tau} \in \mathcal{V}_h^\tau$ be the solution to the SUPG-stabilized time-DG VEM formulation (3.10). We define $u_I^\tau := \Pi_r^t u_I$, and the following error functions:

$$e_I^\tau = u - u_I^\tau, \quad e_\pi^0 = u - \Pi_k^0 u, \quad e_\pi^\nabla = u - \Pi_k^\nabla u, \quad e_{h\tau} = u_I^\tau - u_{h,\tau}.$$

Proposition 5.7 (*A Priori Error Bounds*). *Let Assumption 3.1 and 3.5 hold, then the following bound holds:*

$$\|u - u_{h,\tau}\|_{\mathcal{E}} \le \|e_I^\tau\|_{\mathcal{E}} + \gamma_I \left(\sup_{\substack{v_{h,\tau} \in \mathcal{V}_h^\tau \setminus \{0\}, \\ \|v_{h,\tau}\|_{\mathcal{E}}=1}} \chi_\ell(v_{h,\tau}) + \sup_{\substack{v_{h,\tau} \in \mathcal{V}_h^\tau \setminus \{0\}, \\ \|v_{h,\tau}\|_{\mathcal{E}}=1}} \chi_{\mathrm{supg}}(v_{h,\tau}) + \sup_{\substack{v_{h,\tau} \in \mathcal{V}_h^\tau \setminus \{0\}, \\ \|v_{h,\tau}\|_{\mathcal{E}}=1}} \chi_a(v_{h,\tau}) + \sup_{\substack{v_{h,\tau} \in \mathcal{V}_h^\tau \setminus \{0\}, \\ \|v_{h,\tau}\|_{\mathcal{E}}=1}} \chi_{m,b}(v_{h,\tau}) \right), \tag{5.8}$$

where

$$\chi_\ell(v_{h,\tau}) := \ell_f(v_{h,\tau}) - \ell_f(\Pi_k^{0,K} v_{h,\tau}) + \ell_{u_0}(v_{h,\tau}) - \ell_{u_0}(\Pi_k^{0,K} v_{h,\tau}),$$

$$\chi_{\mathrm{supg}}(v_{h,\tau}) := \ell_f^{\mathrm{supg}}(v_{h,\tau}) - s_{h,\tau}^{\mathrm{supg}}(u_I^\tau, v_{h,\tau}),$$

$$\chi_a(v_{h,\tau}) := \nu a^T(u, v_{h,\tau}) - \nu a_{h,\tau}(u_I^\tau, v_{h,\tau}),$$

$$\chi_{m,b}(v_{h,\tau}) := m^{\partial_t}(u, v_{h,\tau}) + b^T(u, v_{h,\tau}) - m_{h,\tau}^{\partial_t}(u_I^\tau, v_{h,\tau}) - b_{h,\tau}^{\mathrm{skew}}(u_I^\tau, v_{h,\tau}).$$

Proof. The result follows easily from the triangle inequality, the inf–sup stability estimate in Theorem 3.6, the definition of the SUPG-stabilized time-DG VEM in (3.10), and observing that the continuous solution satisfies

$$m^{\partial_t}(u, v_{h,\tau}) + \nu a^T(u, v_{h,\tau}) + b^T(u, v_{h,\tau}) = \ell_f(v_{h,\tau}) + \ell_{u_0}(v_{h,\tau}) \quad \forall v_{h,\tau} \in \mathcal{V}_h^\tau. \quad \square$$

We now estimate each term on the right-hand side of (5.8).

Lemma 5.8 (*Estimate of* $\|\!|e_I^\tau|\!\|_{\mathcal{E}}$). *Under Assumption 3.1 on the mesh-regularity, Assumption 3.5 on the choice of λ_{K_n} and the relation of τ and $h_{\min}$, and Assumption 3.9 on the data of the problem, it holds*

$$
\|\!|e_I^\tau|\!\|_{\mathcal{E}} \lesssim \sum_{n=1}^N \sum_{K \in \Omega_h} \bigg((\tau_n^{2q+1} + \lambda_{K_n} \tau_n^{2q}) \|u\|_{H^{q+1}(I_n; L^2(K))}^2 + \lambda_{K_n} h_K^{2s+2} \|\partial_t u\|_{L^2(I_n; H^{s+1}(K))}^2
$$
$$
+ \bigg(h_K^{2s+2} + \frac{h_K^{2s+2}}{\tau_n} + \nu h_K^{2s} + \lambda_{K_n} \beta_{K_n}^2 h_K^{2s} \bigg) \|u\|_{L^2(I_n; H^{s+1}(K))}^2
$$
$$
+ (\nu + \lambda_{K_n} \beta_{K_n}^2) \tau_n^{2q+2} \|\nabla u\|_{H^{q+1}(I_n; L^2(K)^d)}^2 \bigg). \tag{5.9}
$$

Proof. By the definition of the energy norm $\|\!|\cdot|\!\|_{\mathcal{E}}$ in (3.15), we have

$$\|\!|e_I^\tau|\!\|_{\mathcal{E}}^2 = \|e_I^\tau\|_{L^2(Q_T)}^2 + |e_I^\tau|_J^2 + \nu \|\nabla e_I^\tau\|_{L^2(Q_T)^d}^2 + |e_I^\tau|_{\text{supg}}^2. \tag{5.10}$$

We treat each term on the right-hand side of (5.10) separately.

• *Estimate of* $\|e_I^\tau\|_{L^2(Q_T)}$. Using the triangle inequality, the stability of Π_r^t and its approximation properties in Lemma 5.1, and the VE interpolation estimate in Lemma 5.2, for all $K \in \Omega_h$ and $n \in \{1, \dots, N\}$, we get

$$\|e_I^\tau\|_{L^2(K_n)}^2 \leq 2\|u - \Pi_r^t u\|_{L^2(K_n)}^2 + 2\|\Pi_r^t(u - u_I)\|_{L^2(K_n)}^2 \lesssim \tau_n^{2q+2} \|u\|_{H^{q+1}(I_n; L^2(K))}^2 + h_K^{2s+2} \|u\|_{L^2(I_n; H^{s+1}(K))}^2.$$

• *Estimate of* $|e_I^\tau|_J$. Using the trace inequality in Lemma 5.3, the stability of Π_r^t and its approximation properties in Lemma 5.1, the inverse estimate in (4.3b), and the VE interpolation estimate in Lemma 5.2, we obtain

$$
|e_I^\tau|_J^2 \lesssim \sum_{n=1}^N \bigg(\tau_n^{-1} \|e_I^\tau\|_{L^2(I_n; L^2(\Omega))}^2 + \tau_n \|\partial_t e_I^\tau\|_{L^2(I_n; L^2(\Omega))}^2 \bigg)
$$
$$
\lesssim \sum_{n=1}^N \bigg(\tau_n^{-1} \big(\|u - \Pi_r^t u\|_{L^2(I_n; L^2(\Omega))}^2 + \|\Pi_r^t(u - u_I)\|_{L^2(I_n; L^2(\Omega))}^2 \big) + \tau_n \big(\|\partial_t(u - \Pi_r^t u)\|_{L^2(I_n; L^2(\Omega))}^2 + \|\partial_t(\Pi_r^t(u - u_I))\|_{L^2(I_n; L^2(\Omega))}^2 \big) \bigg)
$$
$$
\lesssim \sum_{n=1}^N \bigg(\tau_n^{2q+1} \|u\|_{H^{q+1}(I_n; L^2(\Omega))}^2 + \sum_{K \in \Omega_h} \frac{h_K^{2s+2}}{\tau_n} \|u\|_{L^2(I_n; H^{s+1}(K))}^2 \bigg).
$$

• *Estimate of* $\nu \|\nabla e_I^\tau\|_{L^2(Q_T)^d}$. Using the triangle inequality, the stability of Π_r^t and its approximation properties in Lemma 5.1, the commutativity of the spatial gradient operator ∇ and the $L^2(\mathcal{T}_\tau)$-orthogonal projection Π_r^t, and the VE interpolation estimate in Lemma 5.2, for all $K \in \Omega_h$ and $n \in \{1, \dots, N\}$, we have

$$\nu \|\nabla e_I^\tau\|_{L^2(K_n)^d}^2 \leq 2\nu \|\nabla u - \Pi_r^t \nabla u\|_{L^2(K_n)^d}^2 + 2\nu \|\Pi_r^t \nabla(u - u_I)\|_{L^2(K_n)^d}^2 \lesssim \nu \tau_n^{2q+2} \|\nabla u\|_{H^{q+1}(I_n; L^2(K)^d)}^2 + \nu h_K^{2s} \|u\|_{L^2(I_n; H^{s+1}(K))}^2.$$

• *Estimate of* $|e_I^\tau|_{\text{supg}}$. We bound this term using the triangle inequality, the commutativity of the first-order time derivative operator ∂_t and the $L^2(K)$-orthogonal projection operator $\Pi_k^{0,K}$, the stability properties of $\Pi_k^{0,K}$, $\Pi_{k-1}^{0,K}$, $\Pi_k^{\nabla,K}$, and Π_r^t, the stability bound in Lemma 5.4, the estimates for Π_r^t in Lemma 5.1, and the VE interpolation estimate in Lemma 5.2, as follows:

$$
|e_I^\tau|_{K_n,\text{supg}}^2 = \lambda_{K_n} \|\partial_t \Pi_k^{0,K} e_I^\tau + \boldsymbol{\beta} \cdot \boldsymbol{\Pi}_{k-1}^{0,K} \nabla e_I^\tau\|_{L^2(K_n)}^2 + \lambda_{K_n} \beta_{K_n}^2 \|\nabla(\text{Id} - \Pi_k^{\nabla,K}) e_I^\tau\|_{L^2(K_n)^d}^2
$$
$$
\lesssim \lambda_{K_n} \|\partial_t e_I^\tau\|_{L^2(K_n)}^2 + \lambda_{K_n} \beta_{K_n}^2 \|\nabla e_I^\tau\|_{L^2(K_n)^d}^2
$$
$$
\lesssim \lambda_{K_n} \big(\|\partial_t(\text{Id} - \Pi_r^t)u\|_{L^2(K_n)}^2 + \|\partial_t \Pi_r^t(u - u_I)\|_{L^2(K_n)}^2 \big)
$$
$$
+ \lambda_{K_n} \beta_{K_n}^2 \big(\|(\text{Id} - \Pi_r^t)\nabla u\|_{L^2(K_n)^d}^2 + \|\Pi_r^t \nabla(u - u_I)\|_{L^2(K_n)^d}^2 \big)
$$
$$
\lesssim \lambda_{K_n} \big(\|\partial_t(\text{Id} - \Pi_r^t)u\|_{L^2(K_n)}^2 + \|\partial_t(u - u_I)\|_{L^2(K_n)}^2 \big)
$$
$$
+ \lambda_{K_n} \beta_{K_n}^2 \big(\|(\text{Id} - \Pi_r^t)\nabla u\|_{L^2(K_n)^d}^2 + \|\nabla(u - u_I)\|_{L^2(K_n)^d}^2 \big)
$$

$$\lesssim \lambda_{K_n} \big(\tau_n^{2q} \|u\|_{H^{q+1}(I_n;L^2(K))}^2 + h_K^{2s+2} \|\partial_t u\|_{L^2(I_n;H^{s+1}(K))}^2 \big)$$
$$+ \lambda_{K_n} \beta_{K_n}^2 \big(\tau_n^{2q+2} \|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}^2 + h_K^{2s} \|u\|_{L^2(I_n;H^{s+1}(K))}^2 \big).$$

Conclusion. Estimate (5.9) follows combining (5.10) with the above four estimates. $\quad\square$

Lemma 5.9 (*Estimate of $\chi_\ell(v_{h,\tau})$*). *Under Assumption 3.1 on the mesh-regularity, Assumption 3.5 on the choice of λ_{K_n} and the relation of τ and $h_{\min}$, and Assumption 3.9 on the data of the problem, for all $v_{h,\tau} \in \mathcal{V}_h^\tau$ with $\|\|v_{h,\tau}\|\|_{\mathcal{E}} = 1$, the term $\chi_\ell(v_{h,\tau})$ can be bounded as follows:*

$$\chi_\ell(v_{h,\tau}) \lesssim \left[\sum_{n=1}^N \sum_{K \in \Omega_h} \min\left\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \right\} h_K^{2s+4} \|f\|_{L^2(I_n;H^{s+1}(K))}^2 + \sum_{K \in \Omega_h} h_K^{2s+2} |u_0|_{H^{s+1}(K)}^2 \right]^{\frac{1}{2}}.$$

Proof. Using the orthogonality properties of $\Pi_k^{0,K}$ and $\Pi_{k-1}^{0,K}$, we have

$$\chi_\ell(v_{h,\tau}) = \ell_f(v_{h,\tau}) - \ell_f(\Pi_k^{0,K} v_{h,\tau}) + \ell_{u_0}(v_{h,\tau}) - \ell_{u_0}(\Pi_k^{0,K} v_{h,\tau})$$
$$= \sum_{n=1}^N \sum_{K \in \Omega_h} \big((\mathrm{Id} - \Pi_k^{0,K})f, (\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau} \big)_{K_n} + \sum_{K \in \Omega_h} \big((\mathrm{Id} - \Pi_k^{0,K})u_0, v_{h,\tau}(\cdot,0) \big)_K$$
$$=: \sum_{n=1}^N \sum_{K \in \Omega_h} \chi_{\mathcal{F}}^{K_n} + \sum_{K \in \Omega_h} \chi_{\mathcal{U}_0}^K. \tag{5.11}$$

Using the scaled Poincaré–Friedrichs inequality in Lemma 5.5, the approximation properties of $\Pi_k^{0,K}$ from Lemma 5.1, and the definition of $|\cdot|_{K_n,\mathrm{supg}}$ in (3.14), for all $K \in \Omega_h$ and $n \in \{1, \dots, N\}$, we get

$$\chi_{\mathcal{F}}^{K_n} = \big((\mathrm{Id} - \Pi_k^{0,K})f, (\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau} \big)_{K_n} \lesssim h_K \|(\mathrm{Id} - \Pi_k^{0,K})f\|_{L^2(K_n)} \|\nabla(\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau}\|_{L^2(K_n)^d}$$
$$\lesssim h_K^{s+2} \|f\|_{L^2(I_n;H^{s+1}(K))} \min\left\{ \frac{|v_{h,\tau}|_{K_n,\mathrm{supg}}}{\lambda_{K_n}^{1/2} \beta_{K_n}}, \|\nabla v_{h,\tau}\|_{L^2(K_n)^d} \right\},$$

which, combined with the Cauchy–Schwarz inequality and the definition of the energy norm $\|\|\cdot\|\|_{\mathcal{E}}$ in (3.15), implies

$$\sum_{n=1}^N \sum_{K \in \Omega_h} \chi_{\mathcal{F}}^{K_n} \lesssim \left(\sum_{n=1}^N \sum_{K \in \Omega_h} h_K^{2s+4} \min\left\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \right\} \|f\|_{L^2(I_n;H^{s+1}(K))}^2 \right)^{\frac{1}{2}} \|\|v_{h,\tau}\|\|_{\mathcal{E}}. \tag{5.12}$$

Finally, using estimate (5.1a) for $\Pi_k^{0,K}$ and the definition of the jump functional $|\cdot|_{\mathsf{J}}$ in (3.13), we obtain

$$\sum_{K \in \Omega_h} \chi_{\mathcal{U}_0}^K \lesssim \left(\sum_{K \in \Omega_h} h_K^{2s+2} |u_0|_{H^{s+1}(K)}^2 \right)^{\frac{1}{2}} \|v_{h,\tau}\|_{L^2(\Sigma_0)} \lesssim \left(\sum_{K \in \Omega_h} h_K^{2s+2} |u_0|_{H^{s+1}(K)}^2 \right)^{\frac{1}{2}} |v_{h,\tau}|_{\mathsf{J}}. \tag{5.13}$$

The desired result follows combining identity (5.11) with estimates (5.12) and (5.13). $\quad\square$

Lemma 5.10 (*Estimate of $\chi_{\mathrm{supg}}(v_{h,\tau})$*). *Under Assumption 3.1 on the mesh-regularity, Assumption 3.5 on the choice of λ_{K_n} and the relation of τ and $h_{\min}$, and Assumption 3.9 on the data of the problem, for all $v_{h,\tau} \in \mathcal{V}_h^\tau$ with $\|\|v_{h,\tau}\|\|_{\mathcal{E}} = 1$, the term $\chi_{\mathrm{supg}}(v_{h,\tau})$ can be bounded as follows:*

$$\chi_{\mathrm{supg}}(v_{h,\tau}) \lesssim \left[\sum_{n=1}^N \sum_{K \in \Omega_h} \left(\frac{\lambda_{K_n} \tau_n^{2q}}{\beta_{K_n}^2} \|u\|_{H^{q+1}(I_n;L^2(K))}^2 + \frac{\lambda_{K_n} h_K^{2s+2}}{\beta_{K_n}^2} \|\partial_t u\|_{L^2(I_n;H^{s+1}(K))}^2 \right. \right.$$
$$+ \left(\frac{\nu^2 \lambda_{K_n} h_K^{2s-2}}{\beta_{K_n}^2} + \lambda_{K_n}(1 + \beta_{K_n}^2) h_K^{2s} \right) \|u\|_{L^2(I_n;H^{s+1}(K))}^2$$
$$+ \left(\frac{\nu^2 \lambda_{K_n} \tau_n^{2q+2}}{h_K^2 \beta_{K_n}^2} + \lambda_{K_n} \tau_n^{2q+2} \right) \|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}^2 + \frac{\lambda_{K_n} h_K^{2s+4}}{\tau_n^2 \beta_{K_n}^2} \|f\|_{L^2(I_n;H^{s+1}(K))}^2$$
$$\left. \left. + \lambda_{K_n}^2 h_K^{2s} \min\left\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \right\} \|\beta\|_{L^\infty(I_n;W^{s,\infty}(K)^d)}^2 \|f\|_{L^2(I_n;H^s(K))}^2 \right) \right]^{\frac{1}{2}}.$$

Proof. Adding and subtracting suitable terms, we get

$$\chi_{\mathrm{supg}}(v_{h,\tau}) = \ell_f^{\mathrm{supg}}(v_{h,\tau}) - s_{h,\tau}^{\mathrm{supg}}(u_I^\tau, v_{h,\tau})$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \Big(\lambda_{K_n}(f, \widetilde{\mathcal{L}}^K v_{h,\tau})_{K_n} - \lambda_{K_n}(\mathcal{L}_{h,\tau}^K u_I^\tau, \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau})_{K_n} - \lambda_{K_n}\beta_{K_n}^2 \int_{I_n} s_a^K((\mathrm{Id} - \Pi_k^{\nabla,K})u_I^\tau, (\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau})\,\mathrm{d}t \Big)$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \Big(\lambda_{K_n}(\mathcal{L}^K u, \widetilde{\mathcal{L}}^K v_{h,\tau})_{K_n} - \lambda_{K_n}(\mathcal{L}_{h,\tau}^K u_I^\tau, \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau})_{K_n} - \lambda_{K_n}\beta_{K_n}^2 \int_{I_n} s_a^K((\mathrm{Id} - \Pi_k^{\nabla,K})u_I^\tau, (\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau})\,\mathrm{d}t \Big)$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \Big(\lambda_{K_n}\big(\mathcal{L}^K u - \mathcal{L}_{h,\tau}^K u_I^\tau, \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau}\big)_{K_n} + \lambda_{K_n}(\mathcal{L}^K u, \partial_t(\mathrm{Id} - \Pi_k^{0,K})v_{h,\tau})_{K_n} + \lambda_{K_n}(\mathcal{L}^K u \boldsymbol{\beta}, (\mathrm{Id} - \Pi_{k-1}^{0,K})\nabla v_{h,\tau})_{K_n}$$

$$- \lambda_{K_n}\beta_{K_n}^2 \int_{I_n} s_a^K((\mathrm{Id} - \Pi_k^{\nabla,K})u_I^\tau, (\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau})\,\mathrm{d}t \Big)$$

$$=: \sum_{n=1}^{N} \sum_{K \in \Omega_h} \big(\chi_{\mathrm{supg},1}^{K_n} + \chi_{\mathrm{supg},2}^{K_n} + \chi_{\mathrm{supg},3}^{K_n} + \chi_{\mathrm{supg},4}^{K_n} \big). \tag{5.14}$$

We bound each term $\chi_{\mathrm{supg},i}$, $i = 1, \ldots, 4$, separately.

Using the Cauchy–Schwarz and the triangle inequalities, and the definition of $|\cdot|_{K_n,\mathrm{supg}}$ in (3.14), we get

$$\chi_{\mathrm{supg},1}^{K_n} = \lambda_{K_n}(\mathcal{L}^K u - \mathcal{L}_{h,\tau}^K u_I^\tau, \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau})_{K_n}$$

$$= \lambda_{K_n}\big(\partial_t(u - \Pi_k^{0,K}u_I^\tau) - \nu\,\mathrm{div}(\nabla u - \Pi_{k-1}^{0,K}\nabla u_I^\tau) + \boldsymbol{\beta}\cdot(\nabla u - \Pi_{k-1}^{0,K}\nabla u_I^\tau), \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau}\big)_{K_n}$$

$$\leq \frac{\lambda_{K_n}^{1/2}}{\beta_{K_n}} \Big(\|\partial_t(u - \Pi_k^{0,K}u_I^\tau)\|_{L^2(K_n)} + \nu\|\mathrm{div}(\nabla u - \Pi_{k-1}^{0,K}\nabla u_I^\tau)\|_{L^2(K_n)} + \|\boldsymbol{\beta}\cdot(\nabla u - \Pi_{k-1}^{0,K}\nabla u_I^\tau)\|_{L^2(K_n)} \Big) |v_{h,\tau}|_{K_n,\mathrm{supg}}. \tag{5.15}$$

We now focus on the local interpolation error terms on the right-hand side of (5.15). Using the triangle inequality, the commutativity of $\Pi_k^{0,K}$ with the first-order time derivative operator ∂_t and that of ∂_t with the VE interpolant operator, the estimates for Π_r^t and $\Pi_k^{0,K}$ in Lemma 5.1, the stability of Π_r^t in Lemma 5.4, and the VE interpolation estimate in Lemma 5.2, it follows that

$$\|\partial_t(u - \Pi_k^{0,K}u_I^\tau)\|_{L^2(K_n)} \leq \|(\mathrm{Id} - \Pi_k^{0,K})\partial_t u\|_{L^2(K_n)} + \|\Pi_k^{0,K}\partial_t(\mathrm{Id} - \Pi_r^t)u\|_{L^2(K_n)} + \|\Pi_k^{0,K}\partial_t\Pi_r^t(u - u_I)\|_{L^2(K_n)}$$

$$\lesssim \|(\mathrm{Id} - \Pi_k^{0,K})\partial_t u\|_{L^2(K_n)} + \|\partial_t(\mathrm{Id} - \Pi_r^t)u\|_{L^2(K_n)} + \|\partial_t(u - u_I)\|_{L^2(K_n)}$$

$$\lesssim h_K^{s+1}\|\partial_t u\|_{L^2(I_n;H^{s+1}(K))} + \tau_n^q\|u\|_{H^{q+1}(I_n;L^2(K))}. \tag{5.16}$$

As for the second term on the right-hand side of (5.15), we use the triangle inequality, the estimates for Π_r^t and $\Pi_{k-1}^{0,K}$ in Lemma 5.1 and their stability properties, the commutativity of Π_r^t and the spatial divergence operator div, and the inverse estimates in (4.3b) and (4.3c) to obtain

$$\nu\|\mathrm{div}(\nabla u - \Pi_{k-1}^{0,K}\nabla u_I^\tau)\|_{L^2(K_n)} \leq \nu\|\mathrm{div}(\nabla u - \Pi_{k-1}^{0,K}\nabla u)\|_{L^2(K_n)} + \nu\|\mathrm{div}\,\Pi_{k-1}^{0,K}(\nabla u - \Pi_r^t\nabla u)\|_{L^2(K_n)} + \nu\|\Pi_r^t\,\mathrm{div}\,\Pi_{k-1}^{0,K}\nabla(u - u_I)\|_{L^2(K_n)}$$

$$\lesssim \nu h_K^{s-1}\|u\|_{L^2(I_n;H^{s+1}(K))} + \nu\frac{\tau^{q+1}}{h_K}\|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}. \tag{5.17}$$

Note that the negative power of h_K could be avoided by asking more space regularity for u, but such effort is not required since this term will be balanced by suitable factors in the final estimates. The last term on the right-hand side of (5.15) can be bounded similarly as follows:

$$\|\boldsymbol{\beta}\cdot(\nabla u - \Pi_{k-1}^{0,K}\nabla u_I^\tau)\|_{L^2(K_n)} \leq \|\boldsymbol{\beta}\cdot(\nabla u - \Pi_r^t\nabla u)\|_{L^2(K_n)} + \|\boldsymbol{\beta}\cdot\Pi_r^t(\nabla u - \Pi_{k-1}^{0,K}\nabla u)\|_{L^2(K_n)} + \|\boldsymbol{\beta}\cdot\Pi_r^t\Pi_{k-1}^{0,K}(\nabla u - \nabla u_I)\|_{L^2(K_n)}$$

$$\lesssim \beta_{K_n}\tau_n^{q+1}\|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)} + \beta_{K_n}h_K^s\|u\|_{L^2(I_n;H^{s+1}(K))}. \tag{5.18}$$

Combining estimates (5.16), (5.17), and (5.18) with identity (5.15), we get

$$\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_{\mathrm{supg},1}^{K_n} \lesssim \Bigg[\sum_{n=1}^{N} \sum_{K \in \Omega_h} \Bigg(\frac{\lambda_{K_n}\tau_n^{2q}}{\beta_{K_n}^2}\|u\|_{H^{q+1}(I_n;L^2(K))}^2 + \frac{\lambda_{K_n}h_K^{2s+2}}{\beta_{K_n}^2}\|\partial_t u\|_{L^2(I_n;H^{s+1}(K))}^2 + \Big(\frac{\nu^2\lambda_{K_n}h_K^{2s-2}}{\beta_{K_n}^2} + \lambda_{K_n}h_K^{2s} \Big)\|u\|_{L^2(I_n;H^{s+1}(K))}^2$$

$$+ \Big(\frac{\nu^2\lambda_{K_n}\tau_n^{2q+2}}{h_K^2\beta_{K_n}^2} + \lambda_{K_n}\tau_n^{2q+2} \Big)\|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}^2 \Bigg) \Bigg]^{\frac{1}{2}}. \tag{5.19}$$

Due to identity (5.6), we get

$$\chi_{\mathrm{supg},2}^{K_n} = \lambda_{K_n}(\mathcal{L}^K u, \partial_t(\mathrm{Id} - \Pi_k^{0,K})v_{h,\tau})_{K_n} = \lambda_{K_n}((\mathrm{Id} - \Pi_k^{0,K})f, (\mathrm{Id} - \Pi_k^{0,K})\partial_t v_{h,\tau})_{K_n},$$

$$\chi_{\mathrm{supg},3}^{K_n} = \lambda_{K_n}(\mathcal{L}^K u \boldsymbol{\beta}, (\mathrm{Id} - \Pi_{k-1}^{0,K})\nabla v_{h,\tau})_{K_n} = \lambda_{K_n}((\mathrm{Id} - \Pi_{k-1}^{0,K})(f\boldsymbol{\beta}), (\mathrm{Id} - \Pi_{k-1}^{0,K})\nabla v_{h,\tau})_{K_n}.$$

Using the polynomial inverse estimate (4.3b), the scaled Poincaré–Friedrichs inequality in Lemma 5.5, the approximation properties in Lemma 5.1 for $\Pi_k^{0,K}$, and the definition in (3.14) of $|\cdot|_{K_n,\mathrm{supg}}$, we have

$$\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_{\mathrm{supg},2}^{K_n} = \sum_{n=1}^{N} \sum_{K \in \Omega_h} \lambda_{K_n}\big((\mathrm{Id} - \Pi_k^{0,K})f, \partial_t(\mathrm{Id} - \Pi_k^{0,K})v_{h,\tau}\big)_{K_n}$$

$$\lesssim \sum_{n=1}^{N} \sum_{K \in \Omega_h} \frac{\lambda_{K_n}}{\tau_n} \|(\mathrm{Id} - \Pi_k^{0,K})f\|_{L^2(K_n)} \|(\mathrm{Id} - \Pi_k^{0,K})v_{h,\tau}\|_{L^2(K_n)}$$

$$\lesssim \sum_{n=1}^{N} \sum_{K \in \Omega_h} \frac{\lambda_{K_n} h_K}{\tau_n} \|(\mathrm{Id} - \Pi_k^{0,K})f\|_{L^2(K_n)} \|\nabla(\mathrm{Id} - \Pi_k^{0,K})v_{h,\tau}\|_{L^2(K_n)^d}$$

$$\lesssim \sum_{n=1}^{N} \sum_{K \in \Omega_h} \frac{\lambda_{K_n}^{1/2} h_K}{\tau_n \beta_{K_n}} \|(\mathrm{Id} - \Pi_k^{0,K})f\|_{L^2(K_n)} |v_{h,\tau}|_{K_n,\mathrm{supg}}$$

$$\lesssim \left(\sum_{n=1}^{N} \sum_{K \in \Omega_h} \frac{\lambda_{K_n} h_K^{2s+4}}{\tau_n^2 \beta_{K_n}^2} \|f\|_{L^2(I_n;H^{s+1}(K))}^2 \right)^{\frac{1}{2}} |v_{h,\tau}|_{\mathrm{supg}}. \tag{5.20}$$

An improved estimate of $\chi_{\mathrm{supg},2}^{K_n}$, valid only for fixed spatial meshes, can be found in Remark 5.11 of [41].

As for $\chi_{\mathrm{supg},3}^{K_n}$, we use the Cauchy–Schwarz inequality, the approximation properties of $\Pi_{k-1}^{0,K}$, the fact that $\nabla \Pi_k^{\nabla,K} v_{h,\tau} \in \mathbb{P}_{k-1}(K)^d$, the definition of $|\cdot|_{K_n,\mathrm{supg}}$ in (3.14), and the Hölder inequality to get

$$\chi_{\mathrm{supg},3}^{K_n} = \lambda_{K_n} \big((\mathrm{Id} - \boldsymbol{\Pi}_{k-1}^{0,K})(f\boldsymbol{\beta}), (\mathrm{Id} - \boldsymbol{\Pi}_{k-1}^{0,K})\nabla v_{h,\tau} \big)_{K_n}$$

$$\lesssim \lambda_{K_n} \|(\mathrm{Id} - \boldsymbol{\Pi}_k^{0,K})(f\boldsymbol{\beta})\|_{L^2(K_n)^d} \|(\mathrm{Id} - \boldsymbol{\Pi}_{k-1}^{0,K})\nabla v_{h,\tau}\|_{L^2(K_n)^d}$$

$$\lesssim \lambda_{K_n} h_K^s \|\boldsymbol{\beta}\|_{L^\infty(I_n;W^{s,\infty}(K)^d)} \|f\|_{L^2(I_n;H^s(K))} \min\left\{ \frac{|v_{h,\tau}|_{K_n,\mathrm{supg}}}{\lambda_{K_n}^{1/2} \beta_{K_n}}, \frac{\|\nabla v_{h,\tau}\|_{L^2(K_n)^d}}{\nu^{1/2}} \right\}.$$

The following estimate can be then obtained using the Cauchy–Schwarz inequality and the definition of the energy norm $\|\|\cdot\|\|_{\mathcal{E}}$:

$$\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_{\mathrm{supg},3}^{K_n} \lesssim \left(\sum_{n=1}^{N} \sum_{K \in \Omega_h} \lambda_{K_n}^2 h_K^{2s} \min\left\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \right\} \|\boldsymbol{\beta}\|_{L^\infty(I_n;W^{s,\infty}(K)^d)}^2 \|f\|_{L^2(I_n;H^s(K))}^2 \right)^{\frac{1}{2}} \|\|v_{h,\tau}\|\|_{\mathcal{E}}. \tag{5.21}$$

Finally, we estimate the term $\chi_{\mathrm{supg},4}^{K_n}$. To do so, we use the stability bound of the bilinear form $s_a^K(\cdot,\cdot)$ in (3.6), the commutativity of Π_r^t and the spatial gradient operator ∇, the stability properties of Π_r^t and $\Pi_k^{\nabla,K}$, the triangle inequality, the definition of $|\cdot|_{K_n,\mathrm{supg}}$ in (3.14), the estimate for $\Pi_k^{\nabla,K}$ in Lemma 5.1, and the VE interpolation estimate in Lemma 5.2 to deduce

$$\chi_{\mathrm{supg},4}^{K_n} = -\lambda_{K_n} \beta_{K_n}^2 \int_{I_n} s_a^K((\mathrm{Id} - \Pi_k^{\nabla,K})u_I^\tau, (\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau})\,\mathrm{d}t$$

$$\leq \hat{\alpha} \lambda_{K_n} \beta_{K_n}^2 \|\Pi_r^t \nabla(\mathrm{Id} - \Pi_k^{\nabla,K})u_I\|_{L^2(K_n)^d} \|\nabla(\mathrm{Id} - \Pi_k^{\nabla,K})v_{h,\tau}\|_{L^2(K_n)^d}$$

$$\leq \hat{\alpha} \lambda_{K_n}^{1/2} \beta_{K_n} \big(\|\nabla(u_I - u)\|_{L^2(K_n)^d} + \|\nabla(\mathrm{Id} - \Pi_k^{\nabla,K})u\|_{L^2(K_n)^d} + \|\nabla \Pi_k^{\nabla,K}(u - u_I)\|_{L^2(K_n)^d} \big) |v_{h,\tau}|_{K_n,\mathrm{supg}}$$

$$\lesssim \lambda_{K_n}^{1/2} \beta_{K_n} h_K^s \|u\|_{L^2(I_n;H^{s+1}(K))} |v_{h,\tau}|_{K_n,\mathrm{supg}}.$$

The sum of the above local estimate over all the elements $K \in \Omega_h$ and the time steps $n \in \{1,\dots,N\}$ leads to

$$\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_{\mathrm{supg},4}^{K_n} \lesssim \left[\sum_{n=1}^{N} \sum_{K \in \Omega_h} \left(\lambda_{K_n} \beta_{K_n}^2 h_K^{2s} \|u\|_{L^2(I_n;H^{s+1}(K))}^2 \right) \right]^{\frac{1}{2}}. \tag{5.22}$$

Combining identity (5.14) with estimates (5.19), (5.20), (5.21), and (5.22), the desired error bound is obtained. $\square$

Lemma 5.11 (*Estimate of* $\chi_a(v_{h,\tau})$). *Under Assumption 3.1 on the mesh-regularity, Assumption 3.5 on the choice of λ_{K_n} and the relation of τ and $h_{\min}$, and Assumption 3.9 on the data of the problem, for all $v_{h,\tau} \in \mathcal{V}_h^\tau$ with $\|\|v_{h,\tau}\|\|_{\mathcal{E}} = 1$, the term $\chi_a(v_{h,\tau})$ can be bounded as follows:*

$$\chi_a(v_{h,\tau}) \lesssim \left(\sum_{n=1}^{N} \sum_{K \in \Omega_h} \left(\nu h_K^{2s} \|u\|_{L^2(I_n;H^{s+1}(K))}^2 + \nu \tau_n^{2q+2} \|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}^2 \right) \right)^{\frac{1}{2}}.$$

Proof. Using the polynomial consistency of the bilinear form $a_{h,\tau}(\cdot,\cdot)$, the stability bound in (3.7), the triangle inequality, the commutativity of Π_r^t and the spatial gradient operator ∇, the stability of Π_r^t, the estimates for $\Pi_k^{\nabla,K}$ in Lemma 5.1 and for u_I in Lemma 5.2, and the Cauchy–Schwarz inequality, we obtain

$$\chi_a(v_{h,\tau}) = \nu a^T(u, v_{h,\tau}) - \nu a_{h,\tau}(u_I^\tau, v_{h,\tau})$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \nu \big((\nabla(\mathrm{Id} - \Pi_k^{\nabla,K})u, \nabla v_{h,\tau})_{K_n} + a_h^{K_n}(\Pi_k^{\nabla,K}u - u_I^\tau, v_{h,\tau}) \big)$$

$$\leq \sum_{n=1}^{N} \sum_{K \in \Omega_h} \nu \big(\|\nabla e_\pi^\nabla\|_{L^2(K_n)^d} + \alpha^* \|\nabla(\Pi_k^{\nabla,K}u - u_I^\tau)\|_{L^2(K_n)^d} \big) \|\nabla v_{h,\tau}\|_{L^2(K_n)^d}$$

$$\leq \sum_{n=1}^{N} \sum_{K \in \Omega_h} \nu \left((1+\alpha^*) \|\nabla e_\pi^\nabla\|_{L^2(K_n)^d} + \alpha^* \|(\mathrm{Id} - \Pi_r^t)\nabla u\|_{L^2(K_n)^d} + \alpha^* \|\Pi_r^t \nabla(u - u_I)\|_{L^2(K_n)^d} \right) \|\nabla v_{h,\tau}\|_{L^2(K_n)^d}$$

$$\lesssim \left(\sum_{n=1}^{N} \sum_{K \in \Omega_h} \left(\nu h_K^{2s} \|u\|_{L^2(I_n; H^{s+1}(K))}^2 + \nu \tau^{2q+2} \|\nabla u\|_{H^{q+1}(I_n; L^2(K)^d)}^2 \right) \right)^{\frac{1}{2}} \||v_{h,\tau}\||_{\mathcal{E}},$$

which completes the proof. $\square$

Lemma 5.12 (*Estimate of $\chi_{m,b}(v_{h,\tau})$*). *Under Assumption 3.1 on the mesh-regularity, Assumption 3.5 on the choice of λ_{K_n} and the relation of τ and $h_{\min}$, and Assumption 3.9 on the data of the problem, for all $v_{h,\tau} \in \mathcal{V}_h^\tau$ with $\||v_{h,\tau}\||_{\mathcal{E}} = 1$, the term $\chi_{m,b}(v_{h,\tau})$ can be bounded as follows:*

$$\chi_{m,b} \lesssim \Bigg(\sum_{n=1}^{N} \sum_{K \in \Omega_h} \Bigg(\Bigg(\Bigg(\|\beta\|_{L^\infty(I_n; W^{1,\infty}(K)^d)}^2 h_K^{2s+2} + \frac{h_K^{2s+2}}{\lambda_{K_n} \beta_{K_n}^2} \|\beta\|_{L^\infty(I_n; W^{s+1}(K)^d)}^2 \Bigg.$$

$$+ \frac{h_K^{2s+2}}{\lambda_{K_n}} + h_K^{2s+2} \min\Bigg\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \Bigg\} \|\beta\|_{L^\infty(I_n; W^{s+1,\infty}(K)^d)}^2$$

$$+ \frac{h_K^{2s+4}}{\lambda_{K_n} \beta_{K_n}^2 \tau_n^2} + \frac{h_K^{2s+4}}{\tau_n^2} \min\Bigg\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \Bigg\} \Bigg) \|u\|_{L^2(I_n; H^{s+1}(K))}^2$$

$$+ \Bigg(\frac{\tau_n^{2q+2}}{\lambda_{K_n}} + h_K^2 \tau_n^{2q} \min\Bigg\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \Bigg\} + \tau_n^{2q+1} + \tau_n^{2q+2} \|\beta\|_{L^\infty(I_n; W^{1,\infty}(K)^d)}^2 \Bigg) \|u\|_{H^{q+1}(I_n; L^2(K))}^2$$

$$+ \Bigg(h_K^2 \tau_n^{2q+2} \|\beta\|_{L^\infty(I_n; W^{1,\infty}(K)^d)}^2 + \frac{h_K^2 \tau_n^{2q+2}}{\lambda_{K_n}} \Bigg) \|\nabla u\|_{H^{q+1}(I_n; L^2(K)^d)}^2 \Bigg) \Bigg)^{\frac{1}{2}}.$$

Proof. Integrating by parts in time, and using the fact that u is continuous in time and the flux–jump identity $\phi(\cdot, t_n^+)\llbracket \psi \rrbracket_n + \psi(\cdot, t_n^-)\llbracket \phi \rrbracket_n = \llbracket \phi \psi \rrbracket_n$, we get the following identity:

$$\chi_{m,b} = \sum_{n=1}^{N} \sum_{K \in \Omega_h} \Bigg(-(u, \partial_t v_{h,\tau})_{K_n} + \int_{I_n} m_h^K(u_I^\tau, \partial_t v_{h,\tau}) + (\beta \cdot \nabla u, v_{h,\tau})_{K_n} \Bigg.$$

$$- \frac{1}{2}(\beta \cdot \Pi_k^{0,K} \nabla u_I^\tau, \Pi_k^{0,K} v_{h,\tau})_{K_n} + \frac{1}{2}(\Pi_k^{0,K} u_I^\tau, \beta \cdot \Pi_k^{0,K} \nabla v_{h,\tau})_{K_n} \Bigg)$$

$$+ \Bigg((u, v_{h,\tau})_{\Sigma_T} - m_h(u_I^\tau(\cdot, T), v_{h,\tau}(\cdot, T)) \Bigg) + \sum_{n=1}^{N-1} \Bigg((u(\cdot, t_n), \llbracket v_{h,\tau} \rrbracket_n)_\Omega - m_h(u_I^\tau(\cdot, t_n^-), \llbracket v_{h,\tau} \rrbracket_n) \Bigg)$$

$$=: \sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_V^{K_n} + \chi_T + \sum_{n=1}^{N-1} \chi_J^n. \tag{5.23}$$

• *Estimate of $\chi_V^{K_n}$.* We first consider the volume terms $\chi_V^{K_n}$. Using the definition of the bilinear form $m_h(\cdot, \cdot)$, the definition of the operators $\widetilde{\mathcal{L}}_{h,\tau}^K$ and $\widetilde{\mathcal{L}}^K$, the skew-symmetry of the bilinear form $b(\cdot, \cdot)$, and the orthogonality properties of $\Pi_k^{0,K}$ and $\boldsymbol{\Pi}_k^{0,K}$, we have

$$\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_V^{K_n} = \sum_{n=1}^{N} \sum_{K \in \Omega_h} \Bigg(-(u, \partial_t v_{h,\tau})_{K_n} + \int_{I_n} m_h^K(u_I^\tau, \partial_t v_{h,\tau}) - (u, \beta \cdot \nabla v_{h,\tau})_{K_n} \Bigg.$$

$$- \frac{1}{2}(\beta \cdot \Pi_k^{0,K} \nabla u_I^\tau, \Pi_k^{0,K} v_{h,\tau})_{K_n} + \frac{1}{2}(\Pi_k^{0,K} u_I^\tau, \beta \cdot \Pi_k^{0,K} \nabla v_{h,\tau})_{K_n} \Bigg)$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \Bigg[-(u, \widetilde{\mathcal{L}}^K v_{h,\tau})_{K_n} + (\Pi_k^{0,K} u_I^\tau, \partial_t \Pi_k^{0,K} v_{h,\tau} + \beta \cdot \boldsymbol{\Pi}_k^{0,K} \nabla v_{h,\tau})_{K_n} \Bigg.$$

$$+ \int_{I_n} s_m^K((\mathrm{Id} - \Pi_k^{0,K}) u_I^\tau, \partial_t(\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau}) \, \mathrm{d}t$$

$$- \frac{1}{2} \Bigg((\beta \cdot \boldsymbol{\Pi}_k^{0,K} \nabla u_I^\tau, \Pi_k^{0,K} v_{h,\tau})_{K_n} + (\Pi_k^{0,K} u_I^\tau, \beta \cdot \boldsymbol{\Pi}_k^{0,K} \nabla v_{h,\tau})_{K_n} \Bigg) \Bigg]$$

$$= \sum_{n=1}^{N} \sum_{K \in \Omega_h} \Bigg[-(u, \partial_t(\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau})_{K_n} - (u\beta, (\mathrm{Id} - \boldsymbol{\Pi}_k^{0,K}) \nabla v_{h,\tau})_{K_n} \Bigg.$$

$$- (\Pi_k^{0,K} u_I^\tau - u, \partial_t \Pi_k^{0,K} v_{h,\tau} + \beta \cdot \boldsymbol{\Pi}_k^{0,K} \nabla v_{h,\tau})_{K_n}$$

$$+ \int_{I_n} s_m^K((\mathrm{Id} - \Pi_k^{0,K}) u_I^\tau, \partial_t(\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau}) \, \mathrm{d}t$$

$$- \frac{1}{2} \Big((\boldsymbol{\beta} \cdot \boldsymbol{\Pi}_k^{0,K} \nabla u_I^\tau, \boldsymbol{\Pi}_k^{0,K} v_{h,\tau})_{K_n} + (\boldsymbol{\Pi}_k^{0,K} u_I^\tau, \boldsymbol{\beta} \cdot \boldsymbol{\Pi}_k^{0,K} \nabla v_{h,\tau})_{K_n} \Big) \Big]$$

$$=: - \sum_{n=1}^N \sum_{K \in \Omega_h} \Big(\chi_{V,1}^{K_n} + \chi_{V,2}^{K_n} + \chi_{V,3}^{K_n} + \chi_{V,4}^{K_n} + \frac{1}{2} \chi_{V,5}^{K_n} \Big).$$

The terms $\chi_{V,1}^{K_n}$ and $\chi_{V,2}^{K_n}$ can be treated similarly to $\chi_{\mathsf{supg},2}$ and $\chi_{\mathsf{supg},3}$ in Lemma 5.10, respectively. The following estimates are then obtained:

$$\sum_{n=1}^N \sum_{K \in \Omega_h} \chi_{V,1}^{K_n} = \sum_{n=1}^N \sum_{K \in \Omega_h} \big(u, \partial_t (\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau} \big)_{K_n}$$

$$\lesssim \Big(\sum_{n=1}^N \sum_{K \in \Omega_h} \frac{h_K^{2s+4}}{\lambda_{K_n} \beta_{K_n}^2 \tau_n^2} \|u\|_{L^2(I_n; H^{s+1}(K))}^2 \Big)^{\frac{1}{2}} |v_{h,\tau}|_{\mathsf{supg}}, \tag{5.24}$$

$$\sum_{n=1}^N \sum_{K \in \Omega_h} \chi_{V,2}^{K_n} = \sum_{n=1}^N \sum_{K \in \Omega_h} (u\boldsymbol{\beta}, (\mathrm{Id} - \Pi_k^{0,K}) \nabla v_{h,\tau})_{K_n}$$

$$\lesssim \Big(\sum_{n=1}^N \sum_{K \in \Omega_h} h_K^{2s+2} \min\Big\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \Big\} \|\boldsymbol{\beta}\|_{L^\infty(I_n; W^{s+1,\infty}(K)^d)}^2 \|u\|_{L^2(I_n; H^{s+1}(K))}^2 \Big)^{\frac{1}{2}} \|\!|v_{h,\tau}|\!\|_\varepsilon. \tag{5.25}$$

The term $\chi_{V,3}^{K_n}$ can be bounded using the definition of $|\cdot|_{K_n,\mathsf{supg}}$ in (3.14), the Cauchy–Schwarz and the triangle inequalities, the estimates for $\Pi_k^{0,K}$ and Π_r^t in Lemma 5.1, and the VE interpolation estimate in Lemma 5.2, as follows:

$$\sum_{n=1}^N \sum_{K \in \Omega_h} \chi_{V,3}^{K_n} = \sum_{n=1}^N \sum_{K \in \Omega_h} \Big((\Pi_k^{0,K} u_I^\tau - u, \widetilde{\mathcal{L}}_{h,\tau}^K v_{h,\tau})_{K_n} + (\Pi_k^{0,K} u_I^\tau - u, \boldsymbol{\beta} \cdot \boldsymbol{\Pi}_k^{0,K}(\mathrm{Id} - \Pi_{k-1}^{0,K}) \nabla v_{h,\tau})_{K_n} \Big)$$

$$\leq \sum_{n=1}^N \sum_{K \in \Omega_h} \lambda_{K_n}^{-1/2} \|\Pi_k^{0,K} u_I^\tau - u\|_{L^2(K_n)} |v_{h,\tau}|_{K_n,\mathsf{supg}}$$

$$\leq \sum_{n=1}^N \sum_{K \in \Omega_h} \lambda_{K_n}^{-1/2} \big(\|\Pi_k^{0,K}(u_I^\tau - u)\|_{L^2(K_n)} + \|e_\pi^0\|_{L^2(K_n)} \big) |v_{h,\tau}|_{K_n,\mathsf{supg}}$$

$$\lesssim \Big(\sum_{n=1}^N \sum_{K \in \Omega_h} \Big(\frac{\tau_n^{2q+2}}{\lambda_{K_n}} \|u\|_{H^{q+1}(I_n; L^2(K))}^2 + \frac{h_K^{2s+2}}{\lambda_{K_n}} \|u\|_{L^2(I_n; H^{s+1}(K))}^2 \Big) \Big)^{\frac{1}{2}} |v_{h,\tau}|_{\mathsf{supg}}. \tag{5.26}$$

As for the term $\chi_{V,4}^{K_n}$, we use the stability bound (3.4), the polynomial inverse estimate (4.3b), the scaled Poincaré–Friedrichs inequality in Lemma 5.5, and similar steps as those used to estimate $\chi_{\mathsf{supg},3}$ in (5.21) to obtain the following estimate:

$$\sum_{n=1}^N \sum_{K \in \Omega_h} \chi_{V,4}^{K_n} = - \sum_{n=1}^N \sum_{K \in \Omega_h} \int_{I_n} s_m^K((\mathrm{Id} - \Pi_k^{0,K}) u_I^\tau, \partial_t (\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau}) \, \mathrm{d}t$$

$$\leq \sum_{n=1}^N \sum_{K \in \Omega_h} \hat{\mu} \|(\mathrm{Id} - \Pi_k^{0,K}) u_I^\tau\|_{L^2(K_n)} \|\partial_t (\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau}\|_{L^2(K_n)}$$

$$\lesssim \sum_{n=1}^N \sum_{K \in \Omega_h} \frac{\mu^* h_K}{\tau_n} \|(\mathrm{Id} - \Pi_k^{0,K}) u_I^\tau\|_{L^2(K_n)} \|\nabla (\mathrm{Id} - \Pi_k^{\nabla,K}) v_{h,\tau}\|_{L^2(K_n)^d}$$

$$\lesssim \sum_{n=1}^N \sum_{K \in \Omega_h} \min\Big\{ \frac{|v_{h,\tau}|_{K_n,\mathsf{supg}}}{\lambda_{K_n}^{1/2} \beta_{K_n}}, \frac{\|\nabla v_{h,\tau}\|_{L^2(K_n)^d}}{\nu^{1/2}} \Big\} \Big(\frac{h_K^{s+2}}{\tau_n} \|u\|_{L^2(I_n; H^{s+1}(K))} + h_K \tau_n^q \|u\|_{H^{q+1}(I_n; L^2(K))} \Big)$$

$$\lesssim \Big(\sum_{n=1}^N \sum_{K \in \Omega_h} \min\Big\{ \frac{1}{\lambda_{K_n} \beta_{K_n}^2}, \frac{1}{\nu} \Big\} \Big(\frac{h_K^{2s+4}}{\tau_n^2} \|u\|_{L^2(I_n; H^{s+1}(K))}^2 + h_K^2 \tau_n^{2q} \|u\|_{H^{q+1}(I_n; L^2(K))}^2 \Big) \Big)^{\frac{1}{2}} \|\!|v_{h,\tau}|\!\|_\varepsilon. \tag{5.27}$$

Adding and subtracting suitable terms, recalling the antisymmetry of the form $b(\cdot, \cdot)$ and using the orthogonality properties of $\Pi_k^{0,K}$ and $\boldsymbol{\Pi}_k^{0,K}$, the following identity can be obtained:

$$\sum_{n=1}^N \sum_{K \in \Omega_h} \chi_{V,5}^{K_n} = \sum_{n=1}^N \sum_{K \in \Omega_h} \Big((\boldsymbol{\beta} \cdot \boldsymbol{\Pi}_k^{0,K} \nabla u_I^\tau, \Pi_k^{0,K} v_{h,\tau})_{K_n} + (\Pi_k^{0,K} u_I^\tau, \boldsymbol{\beta} \cdot \boldsymbol{\Pi}_k^{0,K} \nabla v_{h,\tau})_{K_n} \Big)$$

$$= \sum_{n=1}^N \sum_{K \in \Omega_h} \Big((\boldsymbol{\beta} \cdot (\boldsymbol{\Pi}_k^{0,K} - \mathrm{Id}) \nabla u_I^\tau, \Pi_k^{0,K} v_{h,\tau})_{K_n} + (\boldsymbol{\beta} \cdot \nabla u_I^\tau, v_{h,\tau})_{K_n}$$

$$- (\boldsymbol{\beta} \cdot \nabla u_I^\tau, (\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau})_{K_n} + ((\Pi_k^{0,K} - \mathrm{Id}) u_I^\tau, \boldsymbol{\beta} \cdot \boldsymbol{\Pi}_k^{0,K} \nabla v_{h,\tau})_{K_n}$$

$$- (u_I^\tau, \boldsymbol{\beta} \cdot (\mathrm{Id} - \boldsymbol{\Pi}_k^{0,K}) \nabla v_{h,\tau})_{K_n} + (u_I^\tau, \boldsymbol{\beta} \cdot \nabla v_{h,\tau})_{K_n} \Big)$$

$$
\begin{aligned}
= -\sum_{n=1}^{N}\sum_{K\in\Omega_h} \Big(& ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla u_I^{\tau},(\mathrm{Id}-\boldsymbol{\Pi}_0^{0,K})\boldsymbol{\beta}\,\boldsymbol{\Pi}_k^{0,K}v_{h,\tau})_{K_n} \\
& + ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})(\boldsymbol{\beta}\cdot\nabla u_I^{\tau}),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})v_{h,\tau})_{K_n} \\
& + ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})u_I^{\tau},(\mathrm{Id}-\boldsymbol{\Pi}_0^{0,K})\boldsymbol{\beta}\cdot\boldsymbol{\Pi}_k^{0,K}\nabla v_{h,\tau})_{K_n} \\
& + ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})(u_I^{\tau}\boldsymbol{\beta}),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla v_{h,\tau})_{K_n} \Big) \\
=: -\sum_{n=1}^{N}\sum_{K\in\Omega_h} & (\Theta_1^{K_n}+\Theta_2^{K_n}+\Theta_3^{K_n}+\Theta_4^{K_n}).
\end{aligned}
\tag{5.28}
$$

The first term on the right-hand side of (5.28) can be estimated using the Cauchy–Schwarz inequality, the estimates for $\boldsymbol{\Pi}_k^{0,K}$, $\boldsymbol{\Pi}_0^{0,K}$, and $\boldsymbol{\Pi}_r^{t}$ in Lemma 5.1 and their stability properties, and the VE interpolation estimate in Lemma 5.2, as follows:

$$
\begin{aligned}
\Theta_1^{K_n} &= ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla u_I^{\tau},(\mathrm{Id}-\boldsymbol{\Pi}_0^{0,K})\boldsymbol{\beta}\,\boldsymbol{\Pi}_k^{0,K}v_{h,\tau})_{K_n} \\
&= ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla u,(\mathrm{Id}-\boldsymbol{\Pi}_0^{0,K})\boldsymbol{\beta}\,\boldsymbol{\Pi}_k^{0,K}v_{h,\tau})_{K_n}+((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla(u-u_I^{\tau}),(\mathrm{Id}-\boldsymbol{\Pi}_0^{0,K})\boldsymbol{\beta}\,\boldsymbol{\Pi}_k^{0,K}v_{h,\tau})_{K_n} \\
&\lesssim h_K\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}\Big(\|(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla u\|_{L^2(K_n)^d}+\|\nabla(u-u_I^{\tau})\|_{L^2(K_n)^d} \Big)\|v_{h,\tau}\|_{L^2(K_n)} \\
&\lesssim \|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}\big(h_K^{s+1}\|u\|_{L^2(I_n;H^{s+1}(K))}+h_K\tau_n^{q+1}\|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}\big)\|v_{h,\tau}\|_{L^2(K_n)}.
\end{aligned}
\tag{5.29}
$$

As for the second term on the right-hand side of (5.28), we use the estimates for $\boldsymbol{\Pi}_k^{0,K}$ and $\boldsymbol{\Pi}_r^{t}$ in Lemma 5.1 and their stability properties, the VE interpolation estimate in Lemma 5.2, and the scaled Poincaré–Friedrichs inequality in Lemma 5.5, to get

$$
\begin{aligned}
\Theta_2^{K_n} &= ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})(\boldsymbol{\beta}\cdot\nabla u_I^{\tau}),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})v_{h,\tau})_{K_n} \\
&= ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})(\boldsymbol{\beta}\cdot\nabla u),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})v_{h,\tau})_{K_n}+((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})(\boldsymbol{\beta}\cdot\nabla(u_I^{\tau}-u)),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})v_{h,\tau})_{K_n} \\
&\lesssim h_K\left(h_K^s\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{s,\infty}(K)^d)}\|u\|_{L^2(I_n;H^{s+1}(K))}+\beta_{K_n}\|\nabla(u-u_I^{\tau})\|_{L^2(K_n)^d} \right)\|\nabla(\mathrm{Id}-\boldsymbol{\Pi}_k^{\nabla,K})v_{h,\tau}\|_{L^2(K_n)^d} \\
&\lesssim \left(\frac{h_K^{s+1}}{\lambda_{K_n}^{1/2}\beta_{K_n}}\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{s,\infty}(K)^d)}\|u\|_{L^2(I_n;H^{s+1}(K))}+\frac{h_K\tau_n^{q+1}}{\lambda_{K_n}^{1/2}}\|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}+\frac{h_K^{s+1}}{\lambda_{K_n}^{1/2}}\|u\|_{L^2(I_n;H^{s+1}(K))} \right)|v_{h,\tau}|_{K_n,\mathrm{supg}}.
\end{aligned}
\tag{5.30}
$$

Using the triangle inequality, the estimates for $\boldsymbol{\Pi}_0^{0,K}$, $\boldsymbol{\Pi}_k^{0,K}$, and $\boldsymbol{\Pi}_r^{t}$ in Lemma 5.1 and their stability properties, the VE interpolation estimate in Lemma 5.2, and the VE inverse estimate (4.3a), we obtain the following estimate:

$$
\begin{aligned}
\Theta_3^{K_n} &= ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})u_I^{\tau},(\mathrm{Id}-\boldsymbol{\Pi}_0^{0,K})\boldsymbol{\beta}\cdot\boldsymbol{\Pi}_k^{0,K}\nabla v_{h,\tau})_{K_n} \\
&\lesssim h_K\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}\|(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})u_I^{\tau}\|_{L^2(K_n)}\|\nabla v_{h,\tau}\|_{L^2(K_n)^d} \\
&\lesssim \|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}\left(\|u-u_I^{\tau}\|_{L^2(K_n)}+\|(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})u\|_{L^2(K_n)}+\|\boldsymbol{\Pi}_k^{0,K}(u-u_I^{\tau})\|_{L^2(K_n)} \right)\|v_{h,\tau}\|_{L^2(K_n)} \\
&\lesssim \|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}\left(\tau_n^{q+1}\|u\|_{H^{q+1}(I_n;L^2(K))}+h_K^{s+1}\|u\|_{L^2(I_n;H^{s+1}(K))} \right)\|v_{h,\tau}\|_{L^2(K_n)}.
\end{aligned}
\tag{5.31}
$$

Finally, using the estimates for $\boldsymbol{\Pi}_k^{0,K}$ and $\boldsymbol{\Pi}_r^{t}$ in Lemma 5.1 and their stability properties, the Cauchy–Schwarz inequality, the definition of $|\cdot|_{K_n,\mathrm{supg}}$ in (3.14), and the VE interpolation estimate in Lemma 5.2, we have

$$
\begin{aligned}
\Theta_4^{K_n} &= ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})(u_I^{\tau}\boldsymbol{\beta}),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla v_{h,\tau})_{K_n} \\
&= ((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})(u\boldsymbol{\beta}),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla v_{h,\tau})_{K_n}-((\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})((u-u_I^{\tau})\boldsymbol{\beta}),(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla v_{h,\tau})_{K_n} \\
&\lesssim \left(h_K^{s+1}\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{s+1,\infty}(K)^d)}\|u\|_{L^2(I_n;H^{s+1}(K))}+\beta_{K_n}\|u-u_I^{\tau}\|_{L^2(K_n)} \right)\|(\mathrm{Id}-\boldsymbol{\Pi}_k^{0,K})\nabla v_{h,\tau}\|_{L^2(K_n)^d} \\
&\lesssim \left(\frac{h_K^{s+1}}{\lambda_{K_n}^{1/2}\beta_{K_n}}\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{s+1,\infty}(K)^d)}\|u\|_{L^2(I_n;H^{s+1}(K))}+\frac{\tau_n^{q+1}}{\lambda_{K_n}^{1/2}}\|u\|_{H^{q+1}(I_n;L^2(K))}+\frac{h_K^{s+1}}{\lambda_{K_n}^{1/2}}\|u\|_{L^2(I_n;H^{s+1}(K))} \right)|v_{h,\tau}|_{K_n,\mathrm{supg}}.
\end{aligned}
\tag{5.32}
$$

Combining identity (5.28) with the local estimates (5.29), (5.30), (5.31), and (5.32), we get

$$
\begin{aligned}
\sum_{n=1}^{N}\sum_{K\in\Omega_h}\chi_{V,5}^{K_n} \lesssim \Bigg[\sum_{n=1}^{N}\sum_{K\in\Omega_h} & \bigg(\Big(\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}^2 h_K^{2s+2}+\frac{h_K^{2s+2}}{\lambda_{K_n}\beta_{K_n}^2}\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{s+1,\infty}(K)^d)}^2+\frac{h_K^{2s+2}}{\lambda_{K_n}} \Big)\|u\|_{L^2(I_n;H^{s+1}(K))}^2 \\
& + \Big(h_K^2\tau_n^{2q+2}\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}^2+\frac{h_K^2\tau_n^{2q+2}}{\lambda_{K_n}} \Big)\|\nabla u\|_{H^{q+1}(I_n;L^2(K)^d)}^2 \\
& + \Big(\tau_n^{2q+2}\|\boldsymbol{\beta}\|_{L^{\infty}(I_n;W^{1,\infty}(K)^d)}^2+\frac{\tau_n^{2q+2}}{\lambda_{K_n}} \Big)\|u\|_{H^{q+1}(I_n;L^2(K))}^2 \bigg)\Bigg]^{\frac{1}{2}}\|\!|v_{h,\tau}|\!\|_{\mathcal{E}}.
\end{aligned}
\tag{5.33}
$$

Estimate of χ_T *and* χ_J^n. Using the polynomial consistency of the bilinear form $m_h^K(\cdot, \cdot)$, the stability bound (3.5), the trace inequality in Lemma 5.3, the commutativity of $\Pi_k^{0,K}$ and the VE interpolant with the first-order time derivative operator ∂_t, the Cauchy–Schwarz inequality, and the definition of the upwind-jump functional $|\cdot|_J$ in (3.13), we obtain the following estimate:

$$
\chi_T + \sum_{n=1}^{N-1} \chi_J^n = (u, v_{h,\tau})_{\Sigma_T} - m_h(u_I^\tau(\cdot, T), v_{h,\tau}(\cdot, T)) + \sum_{n=1}^{N-1} \Big((u(\cdot, t_n), [\![v_{h,\tau}]\!]_n)_\Omega - m_h(u_I^\tau(\cdot, t_n^-), [\![v_{h,\tau}]\!]_n) \Big)
$$

$$
= \sum_{K \in \Omega_h} \Big(((\mathrm{Id} - \Pi_k^{0,K})u(\cdot, T), v_{h,\tau}(\cdot, T))_K + m_h^K(\Pi_k^{0,K}u - u_I^\tau(\cdot, T), v_{h,\tau}(\cdot, T)) \Big)
$$

$$
+ \sum_{n=1}^{N-1} \sum_{K \in \Omega_h} \Big(((\mathrm{Id} - \Pi_k^{0,K})u(\cdot, t_n), [\![v_{h,\tau}]\!]_n)_K + m_h^K(\Pi_k^{0,K}u(\cdot, t_n) - u_I^\tau(\cdot, t_n^-), [\![v_{h,\tau}]\!]_n) \Big)
$$

$$
\lesssim \bigg(\sum_{n=1}^N \sum_{K \in \Omega_h} \Big(\tau_n^{-1} \|(\mathrm{Id} - \Pi_k^{0,K})u\|_{L^2(K_n)}^2 + \tau_n \|(\mathrm{Id} - \Pi_k^{0,K})\partial_t u\|_{L^2(K_n)}^2
$$

$$
+ \tau_n^{-1}(\mu^*)^2 \big(\|(\mathrm{Id} - \Pi_k^{0,K})u\|_{L^2(K_n)}^2 + \|u - u_I^\tau\|_{L^2(K_n)}^2 \big)
$$

$$
+ \tau_n(\mu^*)^2 \big(\|(\mathrm{Id} - \Pi_k^{0,K})\partial_t u\|_{L^2(K_n)}^2 + \|\partial_t(u - u_I^\tau)\|_{L^2(K_n)}^2 \big) \Big) \bigg)^{\frac{1}{2}} |v_{h,\tau}|_J
$$

$$
\lesssim \bigg(\sum_{n=1}^N \sum_{K \in \Omega_h} \Big(\frac{h_K^{2s+2}}{\tau_n} \|u\|_{L^2(I_n; H^{s+1}(K))}^2 + \tau_n h_K^{2s+2} \|\partial_t u\|_{L^2(I_n; H^{s+1}(K))}^2 + \tau_n^{2q+1} \|u\|_{H^{q+1}(I_n; L^2(K))}^2 \Big) \bigg)^{\frac{1}{2}} |v_{h,\tau}|_J. \tag{5.34}
$$

The desired result is then obtained by combining identity (5.23) with estimates (5.24), (5.25), (5.26), (5.27), (5.33), and (5.34). □

Proof of Theorem 3.10. Combining the *a priori* error bound in Proposition 5.7 with the estimates in Lemmas 5.8, 5.9, 5.10, 5.11, and 5.12 one obtains a *general convergence result*, which underlines the different local contributions. The proof of Theorem 3.10 follows as a simplified case, simply by elaborating on the bounds using the specific theorem assumptions and finally dropping all the regularity terms.

5.3. Avoiding degeneration of the error estimates for $\tau \ll h$

Under suitable conditions, we are able to eliminate the terms where τ appears at the denominator in Theorem 3.10. In the present section, we describe briefly the involved modifications. As a starting point, we obviously set ourselves within the range of assumptions outlined in Theorem 3.10.

Differently from the previous part, see Remark 3.4, we now require that the mesh Ω_h is fixed and does not change from one time-slab to the next. This is quite natural as the "transfer" error induced by a change of mesh depends on the spatial meshsize h, but the number of such occurrences grows as τ^{-1}. Furthermore, we require the polynomial order in time $r \geq 1$. Finally, we assume (mainly for simplicity of exposition) that the spatial mesh (family) is quasi-uniform.

The first key point is substituting the space–time projection operator introduced at the beginning of Section 5.2 with the following approximant, for which we keep the same notation. Given $u \in H^1(0, T; L^2(\Omega))$ the discrete function $u_I^\tau \in \mathcal{V}_h^\tau$ is *continuous in time* and defined by

$$
\begin{cases}
m_h(u_I^\tau(\cdot, t_n), v_h) = (u(\cdot, t_n), v_h)_{0,\Omega} & \forall n \in \{0, 1, \ldots, N\}, \ \forall v_h \in \mathcal{V}_h, \\
\displaystyle\int_{t_{n-1}}^{t_n} m_h(\partial_t u_I^\tau(\cdot, t), (t - t_{n-1})^l v_h) \, dt = \int_{t_{n-1}}^{t_n} (\partial_t u(\cdot, t), (t - t_{n-1})^l v_h) \, dt \\
\qquad\qquad \forall n \in \{1, \ldots, N\}, \ \forall v_h \in \mathcal{V}_h, \forall l \in \{1, 2, \ldots, r-1\}.
\end{cases}
$$

It can be checked that the above interpolant satisfies, for all $n = 1, 2, \ldots, N$, the continuity properties

$$
\|u_I^\tau\|_{L^2(I_n, L^2(\Omega))} \lesssim \|u\|_{L^2(I_n, L^2(\Omega))} + \tau_n \|\partial_t u\|_{L^2(I_n, L^2(\Omega))} \quad \text{and} \quad \|\partial_t u_I^\tau\|_{L^2(I_n, L^2(\Omega))} \lesssim \|\partial_t u\|_{L^2(I_n, L^2(\Omega))},
$$

and also that optimal (in h, τ) approximation error bounds hold for $(u - u_I^\tau)$ in the various norms of interest needed in our analysis.

Thanks to the continuity of u_I^τ, the error term $|e_I^\tau|_J^2$ in Lemma 5.8 vanishes; note that such quantity was the only responsible for the $1/\tau_n$ term appearing in Lemma 5.8. Furthermore, thanks to the peculiar definition of u_I^τ here above, the term $\chi_T + \sum_{n=1}^{N-1} \chi_J^n$ appearing in (5.34) also vanishes. In the simpler finite element case, which is included in our analysis as explained in Section 3.4, these are the only terms leading to negative powers of τ in our error estimate.

In the more general case of virtual elements, we need to modify the scheme, adding the following additional stabilization term to the discrete form $\mathcal{B}_{h,\tau}^{\mathrm{supg}}(\cdot, \cdot)$:

$$
\sum_{K \in \Omega_h} \sum_{n=1}^N \lambda_{K_n} \int_{I_n} s_m^K(\partial_t(\mathrm{Id} - \Pi_k^{0,K})u_{h,\tau}, \partial_t(\mathrm{Id} - \Pi_k^{0,K})v_{h,\tau}) \, dt, \tag{5.35}
$$

which, quite trivially, leads to a stability bound controlling a stronger norm $\|\|u_{h,\tau}\|\|_{\mathcal{E}}^2$, c.f. (3.15), now including also the term

$$\sum_{K \in \Omega_h} \sum_{n=1}^{N} \lambda_{K_n} \|\partial_t (\mathrm{Id} - \Pi_k^{0,K}) v_{h,\tau}\|_{L^2(K_n)}^2. \tag{5.36}$$

The control on the test function in the above norm allows us to deal with all the remaining "bad" terms in our convergence analysis, namely $\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_{\mathrm{supg},2}^{K_n}$ in Lemma 5.10, plus $\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_{V,1}^{K_n}$ and $\sum_{n=1}^{N} \sum_{K \in \Omega_h} \chi_{V,4}^{K_n}$ in Lemma 5.12. Indeed, we can now avoid using the inverse estimate in time and simply exploit directly control on (5.36); the Cauchy–Schwarz inequality and standard manipulations yield the error terms

- $\sum_{n=1}^{N} \sum_{K \in \Omega_h} \lambda_{K_n} \|(\mathrm{Id} - \Pi_k^{0,K}) f\|_{L^2(K_n)}^2 \lesssim h^{2k+3},$

- $\sum_{n=1}^{N} \sum_{K \in \Omega_h} \lambda_{K_n}^{-1} \|(\mathrm{Id} - \Pi_k^{0,K}) u\|_{L^2(K_n)}^2 \lesssim \max\{\nu, h\}\, h^{2k},$

- $\sum_{n=1}^{N} \sum_{K \in \Omega_h} \int_{I_n} \lambda_{K_n}^{-1} s_m^K ((\mathrm{Id} - \Pi_k^{0,K}) u_I^{\tau}, (\mathrm{Id} - \Pi_k^{0,K}) u_I^{\tau}) \lesssim \max\{\nu, h\}\, \left(h^{2k} + \tau^{2k}(\tau/h)^2 \right).$

The bounds here above can be easily obtained by the same techniques used in the rest of this contribution and therefore we avoid showing the details. Finally, it can be checked that term (5.35) is also of optimal order with respect to the interpolation error, as usual assuming sufficient regularity of the solution u.

Remark 5.13 (*The Case $r = 0$*). In the case $r = 0$, we cannot take a continuous-in-time interpolant, but we can choose the unique approximant that satisfies

$$m_h(u_I^{\tau}(\cdot, t_n), v_h) = (u(\cdot, t_n^-), v_h)_{0,\Omega} \qquad \forall n \in \{1, 2, \dots, N\}, \ \forall v_h \in \mathcal{V}_h.$$

It is immediate to check that this choice is still sufficient to make term (5.34) vanish. In order to deal with $|e_I^{\tau}|_J^2$ in Lemma 5.8, we simply avoid such term in the interpolation estimates. Therefore, in order to avoid negative powers of τ, the final error bound for the $r = 0$ case will be in the weaker norm

$$\|\|w_h\|\|_{\mathcal{E},r=0}^2 := \|w_h\|_{L^2(\Sigma_T)}^2 + \|w_h\|_{L^2(Q_T)}^2 + \nu \|\nabla w_h\|_{L^2(Q_T)^d}^2 + |w_h|_{\mathrm{supg}}^2$$

where the jump terms have been excluded. Finally, note that in the $r = 0$ case there is no need to introduce (5.35) since all the associated terms are now vanishing. ∎

6. Numerical tests

In this section, we present some numerical results in three space dimensions in order to validate the theoretical derivations from the practical perspective, and evidence the effectiveness of the proposed stabilized scheme compared to a non-stabilized approach. For the sake of efficiency, we make use of the serendipity version of VEM, see Section 3.4. We will refer to the stabilized method, i.e., the scheme described in (3.10) as SUPG, whereas the scheme *without* the SUPG stabilization terms will be denoted as NONE. As described in the previous sections, the SUPG scheme depends on a set of parameters. In all the numerical experiments of this section, we will set these parameters as

$$\zeta = 0.1, \quad C_{\mathrm{inv}} \approx 10k^2 \quad \text{and} \quad \lambda_{K_n} = \zeta \min\left\{ \frac{h_K^2}{\nu C_{\mathrm{inv}}^2}, \frac{h_K}{\bar{\beta}_{Q_T}} \right\}.$$

An (interpolatory) Lagrangian basis is used for the space $\mathbb{P}_k(I_n)$ for $n = 1, \dots, N$, so that the associated degrees of freedom (DoFs) are pointwise evaluations. Finally, we use the classical dofi-dofi choice (see, e.g., [17, §4.6]) for the VE stabilization term.

The numerical experiments are organized as follows. In Section 6.1, we evaluate the convergence trend of the schemes SUPG and NONE, both in the convection- and the diffusion-dominated regimes. Then, in Section 6.2, we apply these two schemes to a benchmark problem similar to the "Three body movement" proposed in [7, §5.1].

6.1. A standard convergence test

In this section, we solve the time-dependent convection–diffusion problem in the space–time domain $Q_T = (0, 1)^3 \times (0, 1.5)$. As transport advective field, we set the following space–time dependent function:

$$\boldsymbol{\beta}(x, y, z, t) := \begin{bmatrix} e^{t/2} \sin(\pi(x + y + 2z)) \\ e^{t/2} \sin(\pi(x + y + 2z)) \\ -e^{t/2} \sin(\pi(x + y + 2z)) \end{bmatrix}$$

that increases exponentially with time. We consider the following two values of the diffusive coefficients:

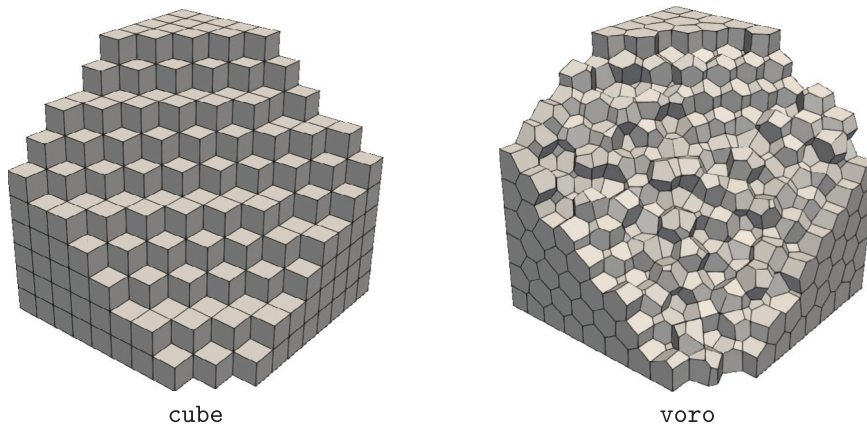$$\nu = 1 \quad \text{and} \quad \nu = 10^{-10},$$

cube   voro

Fig. 1. The interior of some spatial meshes of each type with approximately the same meshsize.

so as to obtain a problem characterized by a diffusion- or a convection-dominated regime, respectively. Then, the right-hand side will be properly modified according ν so that the exact solution is given by

$$u(x,\,y,\,z,\,t) = e^{0.3t} \sin(\pi x) \cos(\pi y) \sin(\pi z).$$

In the proposed tests, we consider a family of four spatial meshes with decreasing meshsize h, while the time domain is split in uniform intervals with time step $\tau \simeq h$. In all the numerical experiments, the VE approximation degree coincides with the time polynomial degree, and we refer to it simply as k. Moreover, to test the robustness of the proposed method with respect to element distortion, we consider two different types of spatial meshes:

- `cube`: structured meshes composed by cubes, see Fig. 1(left panel);
- `voro`: meshes composed by polyhedral elements that may have small edges or faces, see Fig. 1(right panel).

For both schemes, we compute the following errors:

- the H^1 **seminorm error** at the final time

$$e_{H^1}^{T} := \Big(\sum_{K \in \Omega_h} \|\nabla(u - \Pi_k^{\nabla} u_{h,\tau})(\cdot, T)\|_{L^2(K)^3}^2 \Big)^{\frac{1}{2}};$$

- the L^2 **norm error** at final time

$$e_{L^2}^{T} := \Big(\sum_{K \in \Omega_h} \|(u - \Pi_k^0 u_{h,\tau})(\cdot, T)\|_{L^2(K)}^2 \Big)^{\frac{1}{2}};$$

- the H^1 **norm error** on the space–time cylinder

$$e_{H^1}^{Q_T} := \Big(\sum_{n=1}^{N} \sum_{K \in \Omega_h} \big(\|u - \Pi_k^0 u\|_{L^2(K_n)}^2 + \|\nabla(u - \Pi_k^{\nabla} u)\|_{L^2(K_n)^3}^2 \big) \Big)^{\frac{1}{2}}.$$

For the scheme SUPG, under the above condition $\tau \simeq h$, the following asymptotic behaviour is expected:

$$e_{H^1}^{T} = \mathcal{O}\left(h^k\right), \qquad e_{L^2}^{T} = \mathcal{O}\left(h^{k+1}\right), \qquad e_{H^1}^{Q_T} = \mathcal{O}\left(h^k\right),$$

in both the convection- and the diffusion-dominated regimes. Furthermore, only for the solution $u_{h,\tau}$ obtained with the scheme SUPG, we compute also the following quantity:

$$e_{\mathcal{V}_h^\tau} := \|\|u_{h,\tau} - u_I\|\|_{\mathcal{E}},$$

where u_I is the DoF-interpolant of the exact solution u. According to Theorem 3.10, the asymptotic behaviour of $e_{\mathcal{V}_h^\tau}$ depends on the regime we are considering. More specifically, in the convection-dominated regime, it decays as $\mathcal{O}(h^{k+1/2})$, whereas, in a diffusion dominated regime, it decays as $\mathcal{O}(h^k)$.

In Fig. 2, we show the errors obtained for $k = 1$ and 2 in the diffusion-dominated regime, i.e., for $\nu = 1$. The schemes SUPG and NONE have the expected convergence rates. Moreover, for the same mesh and approximation degree, the absolute values of the errors obtained with SUPG and NONE are very close to each other. This is a numerical evidence that the stabilization terms added
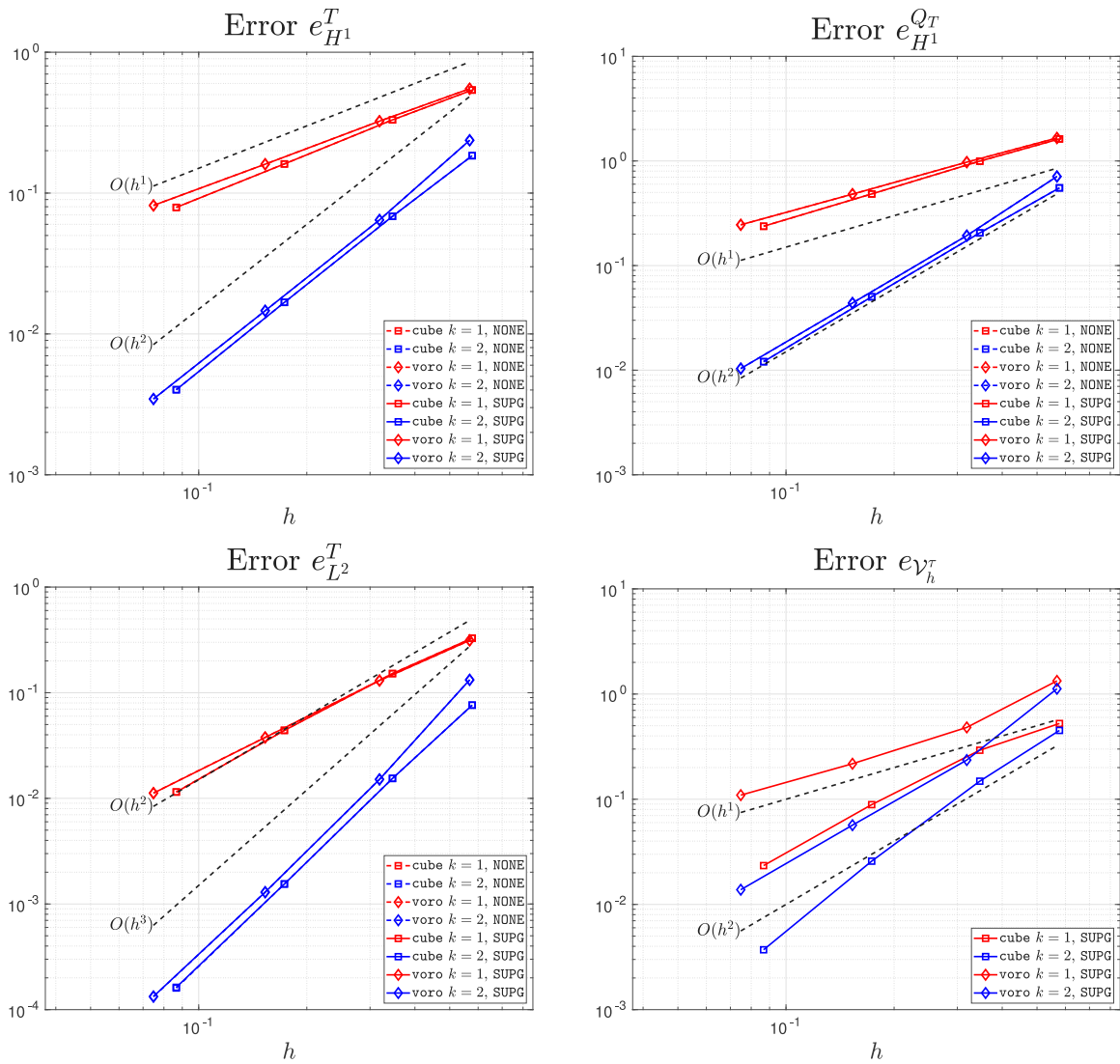
Fig. 2. Convergence Analysis: The trend of all errors taken into account in the diffusion-dominated regime, i.e., for $v = 1$.

in the SUPG scheme do not affect the convergence rates in the diffusion-dominated regime. In Fig. 2(fourth panel), we observe a superconvergence trend of the error $e_{\mathcal{V}_h^\tau}$ for cube meshes and $k = 1$; this may be due to the shape regularity of the mesh and the fact that we are evaluating an error based only on DoF values.

Now, we consider the convection-dominated regime. In Fig. 3, we show the errors obtained for $v = 10^{-10}$. In this case, the results for the two approximation degrees considered are different. Indeed, for $k = 1$, a similar behaviour is observed for the schemes SUPG and NONE. More precisely, the convergence lines are close to each other for the errors $e_{H^1}^T$ and $e_{H^1}^{Q_T}$, whereas, for the error $e_{L^2}^T$, the error trend for the NONE scheme is not optimal in the last refinement step. The advantage of using the SUPG scheme becomes more evident for $k = 2$. For *all* errors and for *both* type of meshes, the convergence rates for the NONE scheme degrade. Moreover, if we compare the error $e_{\mathcal{V}_h^\tau}$ in Fig. 2 and 3, we obtain an additional $1/2$ in the convergence rates in the convection-dominated regime, which is in agreement with Theorem 3.10. This fact is more evident for the voro meshes, which do not satisfy any shape regularity that may affect the trend of the error.

6.2. A qualitative purely advective test

In this section, we make a qualitative assessment of the proposed scheme. To achieve this goal, we produce a benchmark problem similar to the bi-dimensional Example 2 of [7, §5.1], here developed in three space dimensions. In that work, the authors considered three disjoint bodies subject to a rotating advection field and set the diffusive coefficient to 10^{-20} to mimic a transport problem.
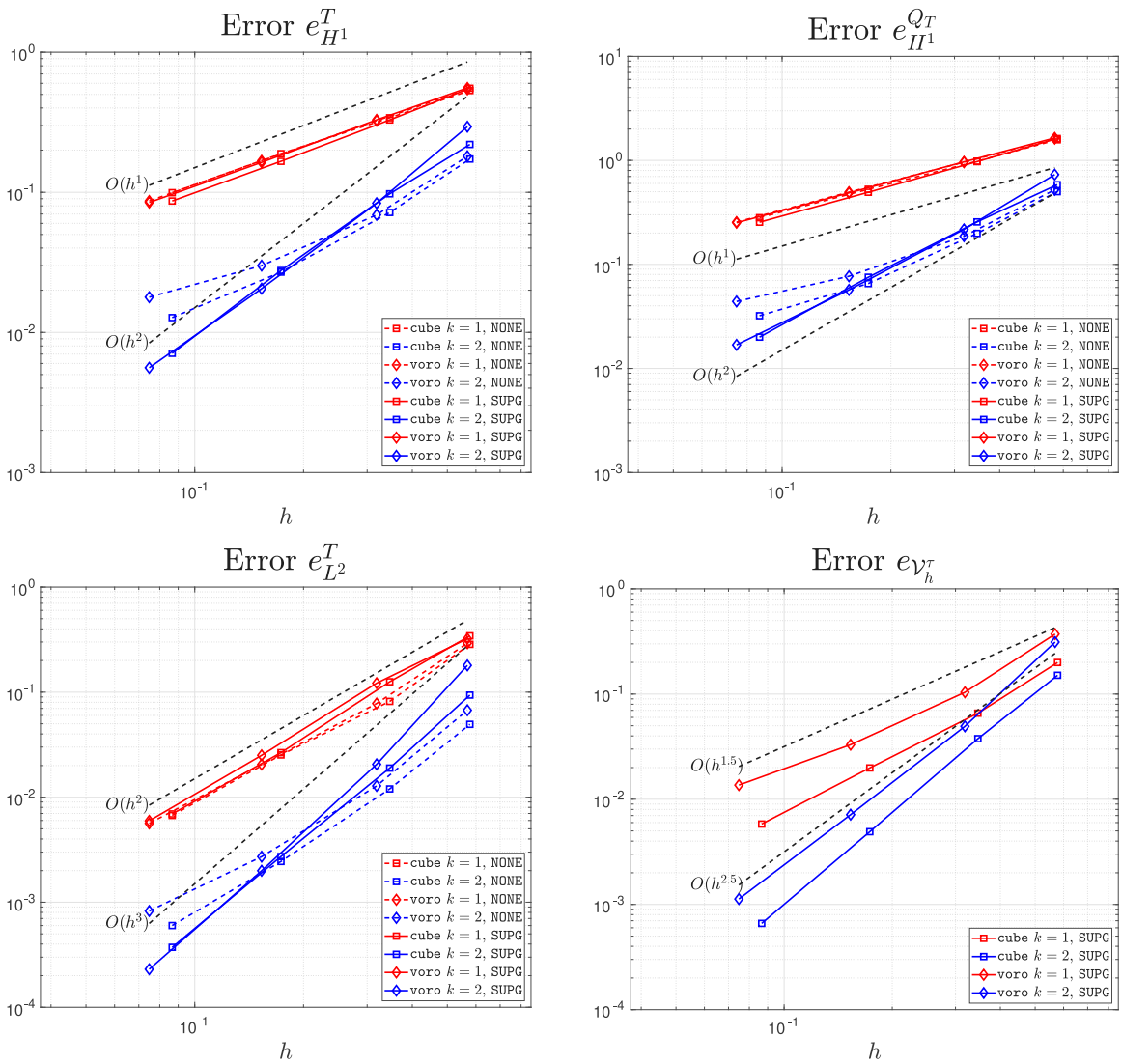
Fig. 3. Convergence Analysis: The trend of all errors taken into account in the convection-dominated regime, i.e., for $\nu = 10^{-10}$.

Let $\Omega = (0, 1)^3$ and the initial condition be given by

$$u_0(x, y, z) := \begin{cases} 1 & \text{if } \sqrt{(x - 0.25)^2 + (y - 0.50)^2 + (z - 0.50)^2} \leq 0.2 \\ 0 & \text{otherwise} \end{cases}$$

that represents a ball of radius $r = 0.2$ centred in $C(0.25, 0.5, 0.5)$. We set $f = 0$, $\nu = 10^{-20}$, and the advection field

$$\beta(x, y, z, t) := \begin{bmatrix} 0.5 - y \\ x - 0.5 \\ 0.0 \end{bmatrix}.$$

For these data, the ball is expected to rotate around the barycentre of the unit cube and, since the diffusive coefficient is close to zero, it has to preserve its shape.

We compute the discrete solution for $k = 1$, a fine tetrahedral spatial mesh, a fixed time step $\tau = 10^{-1}$, and a final time $T = 6$.

In Fig. 4, we show some clips of the discrete solution at different times for both the NONE and the SUPG schemes, respectively. Such clips are obtained using the "clip" filter of Paraview [42] where we associate to each mesh vertex the value of the discrete function $u_{h,\tau}$. As a consequence, we do not see exact circles in such plots, as the shape is affected by the aforementioned geometric interpolation and visualization procedure.

At $t = 0$, we plot the initial condition for both cases. For *all* the other time instances shown in Fig. 4, several spurious values of $u_{h,\tau}$ appear in the interior of the domain Ω with the NONE scheme, i.e., the discrete solution $u_{h,\tau}$ is not constantly zero outside
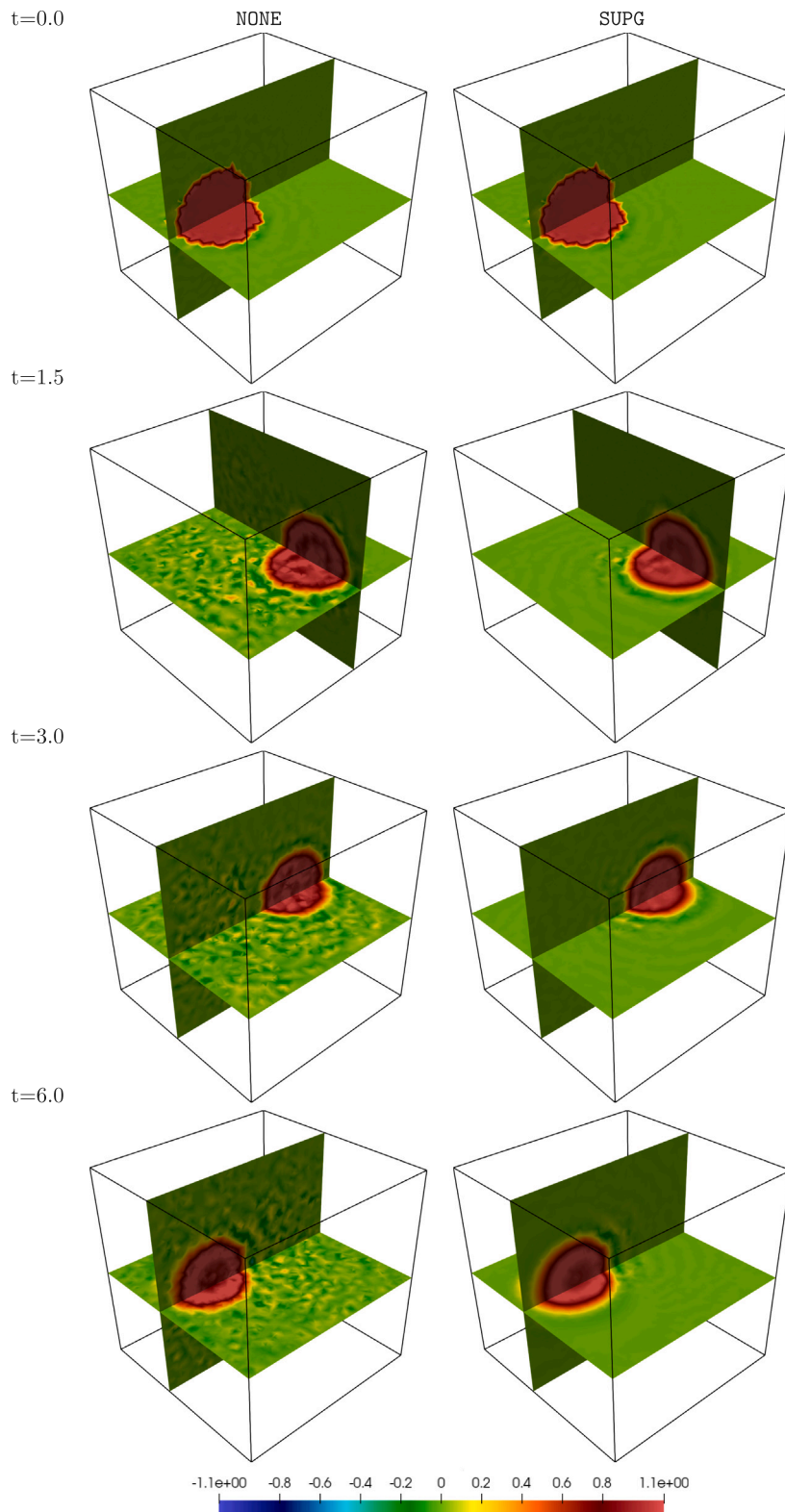
Fig. 4. Benchmark problem: Comparison between the discrete solutions obtained with the schemes NONE and SUPG at different time steps.

the moving sphere. Such instabilities increase with time, as observed in the results obtained at the final time $T = 6$.

The SUPG scheme does not exhibit these instabilities, as the solution outside the sphere is more uniform and, according to the colorbar, it is closer to zero then the NONE discrete solution. Furthermore, despite the effect of the Paraview interpolation, the sphere seems more uniform and rounded.

This benchmark test highlights the importance of using a stabilized scheme even for $k = 1$, a point that was not evident from the error plots of Section 6.1, c.f. the errors of NONE and SUPG for $k = 1$ in Figs. 2 and 3.

7. Conclusions

In this work, we considered a high-order SUPG-stabilized fully discrete scheme that combines finite or virtual element spatial discretizations with an upwind-DG time-stepping. For this fully discrete scheme with finite element spatial discretizations, a robust analysis was missing in the literature. Moreover, this is the first work where a SUPG stabilization has been considered in a high-order-in-time fully discrete setting with virtual element spatial discretizations.

Using nonstandard test functions, we have shown that the method is inf–sup stable with respect to a norm involving an $L^2(0, T; L^2(\Omega))$-term without requiring any transformation of the original problem. Such a stability estimate is used to show that the method is robust and provides optimal convergence rates in the convection- and diffusion-dominated regimes.

We have presented some numerical experiments in $(3 + 1)$-dimensions that show the robustness of the method, as well as the expected convergence rates of order $\mathcal{O}(h^{k+\frac{1}{2}})$ for the error in the energy norm, in the convection-dominated regime $0 < \nu \ll 1$.

CRediT authorship contribution statement

L. Beirão da Veiga: Writing – original draft, Software, Formal analysis, Conceptualization. **F. Dassi:** Writing – original draft, Visualization, Software, Conceptualization. **S. Gómez:** Writing – original draft, Software, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Data availability

No data was used for the research described in the article.

References

[1] F. Shakib, T.J.R. Hughes, A new finite element formulation for computational fluid dynamics. IX. Fourier analysis of space–time Galerkin/least-squares algorithms, Comput. Methods Appl. Mech. Engrg. 87 (1) (1991) 35–58.

[2] A.N. Brooks, T.J.R. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg. 32 (1982) 199–259.

[3] C. Johnson, J. Saranen, Streamline diffusion methods for the incompressible Euler and Navier- Stokes equations, Math. Comp. 47 (1986) 1–18.

[4] L.P. Franca, S.L. Frey, T.J.R. Hughes, Stabilized finite element methods. I.: Application to the advective- diffusive model, Comput. Methods Appl. Mech. Engrg. 95 (2) (1992) 253–276.

[5] P.B. Bochev, M.D. Gunzburger, J.N. Shadid, Stability of the SUPG finite element method for transient advection-diffusion problems, Comput. Methods Appl. Mech. Engrg. 193 (23–26) (2004) 2301–2323.

[6] V. John, E. Schmeyer, Finite element methods for time-dependent convection–diffusion-reaction equations with small diffusion, Comput. Methods Appl. Mech. Engrg. 198 (3–4) (2008) 475–494.

[7] N. Ahmed, G. Matthies, L. Tobiska, H. Xie, Discontinuous Galerkin time stepping with local projection stabilization for transient convection–diffusion-reaction problems, Comput. Methods Appl. Mech. Engrg. 200 (21–22) (2011) 1747–1756.

[8] Z. Dong, H. Li, A space–time finite element method based on local projection stabilization in space and discontinuous Galerkin method in time for convection–diffusion-reaction equations, Appl. Math. Comput. 397 (2021) 15, Id/No 125937.

[9] E. Burman, M.A. Fernández, Finite element methods with symmetric stabilization for the transient convection–diffusion-reaction equation, Comput. Methods Appl. Mech. Engrg. 198 (33–36) (2009) 2508–2519.

[10] A. Ern, F. Schieweck, Discontinuous Galerkin method in time combined with a stabilized finite element method in space for linear first-order PDEs, Math. Comp. 85 (301) (2016) 2099–2129.

[11] E. Burman, G. Smith, Analysis of the space semi-discretized SUPG method for transient convection–diffusion equations, Math. Models Methods Appl. Sci. 21 (10) (2011) 2049–2068.

[12] V. John, J. Novo, Error analysis of the SUPG finite element discretization of evolutionary convection–diffusion-reaction equations, SIAM J. Numer. Anal. 49 (3) (2011) 1149–1176.

[13] J. de Frutos, B. García-Archilla, J. Novo, Local error estimates for the SUPG method applied to evolutionary convection-reaction–diffusion equations, J. Sci. Comput. 66 (2) (2016) 528–554.

[14] J. Zhang, X. Liu, Uniform stability of the SUPG method for the evolutionary convection–diffusion-reaction equation, Comput. Math. Appl. 124 (2022) 1–6.

[15] N. Ahmed, G. Matthies, Higher order continuous Galerkin-Petrov time stepping schemes for transient convection–diffusion-reaction equations, ESAIM Math. Model. Numer. Anal. 49 (5) (2015) 1429–1450.

[16] S. Srivastava, S. Ganesan, Local projection stabilization with discontinuous Galerkin method in time applied to convection dominated problems in time-dependent domains, BIT 60 (2) (2020) 481–507.

[17] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L.D. Marini, A. Russo, Basic principles of virtual element methods, Math. Models Methods Appl. Sci. 23 (1) (2013) 199–214.

[18] M.F. Benedetto, S. Berrone, A. Borio, S. Pieraccini, S. Scialò, Order preserving SUPG stabilization for the virtual element formulation of advection-diffusion problems, Comput. Methods Appl. Mech. Engrg. 311 (2016) 18–40.

[19] L. Beirão da Veiga, F. Dassi, C. Lovadina, G. Vacca, SUPG-stabilized virtual elements for diffusion-convection problems: a robustness analysis, ESAIM Math. Model. Numer. Anal. 55 (5) (2021) 2233–2258.

[20] Y. Li, M. Feng, A local projection stabilization virtual element method for convection–diffusion-reaction equation, Appl. Math. Comput. 411 (2021) 15, Paper No. 126536.

[21] L. Beirão da Veiga, C. Lovadina, M. Trezzi, CIP-stabilized virtual elements for diffusion-convection-reaction problems, IMA J. Numer. Anal. (2024) http://dx.doi.org/10.1093/imanum/drae020.

[22] N. Ahmed, G. Matthies, Numerical study of SUPG and LPS methods combined with higher order variational time discretization schemes applied to time-dependent linear convection–diffusion-reaction equations, J. Sci. Comput. 67 (3) (2016) 988–1018.

[23] W. McLean, Strongly Elliptic Systems and Boundary Integral Equations, Cambridge University Press, Cambridge, 2000.

[24] L.C. Evans, Partial Differential Equations, second ed., in: Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010.

[25] L. Beirão da Veiga, C. Lovadina, A. Russo, Stability analysis for the virtual element method, Math. Models Methods Appl. Sci. 27 (13) (2017) 2557–2594.

[26] S.C. Brenner, L.-Y. Sung, Virtual element methods on meshes with small edges or faces, Math. Models Methods Appl. Sci. 28 (7) (2018) 1291–1336.

[27] S. Cao, L. Chen, Anisotropic error estimates of the linear virtual element method on polygonal meshes, SIAM J. Numer. Anal. 56 (5) (2018) 2913–2939.

[28] B. Ahmad, A. Alsaedi, F. Brezzi, L.D. Marini, A. Russo, Equivalent projectors for virtual element methods, Comput. Math. Appl. 66 (3) (2013) 376–391.

[29] L. Beirão da Veiga, F. Dassi, A. Russo, High-order virtual element method on polyhedral meshes, Comput. Math. Appl. 74 (5) (2017) 1110–1122.

[30] L. Mascotto, The role of stabilization in the virtual element method: a survey, Comput. Math. Appl. 151 (2023) 244–251.

[31] L. Beirão da Veiga, F. Brezzi, L.D. Marini, A. Russo, Serendipity nodal VEM spaces, Comput. & Fluids 141 (2016) 2–12.

[32] L. Beirão da Veiga, L. Mascotto, Stability and interpolation properties of serendipity nodal virtual elements, Appl. Math. Lett. 142 (2023) 9, Paper No. 108639.

[33] G. Vacca, An H^1-conforming virtual element for Darcy and Brinkman equations, Math. Models Methods Appl. Sci. 28 (1) (2018) 159–194.

[34] J. Huang, Y. Yu, Some estimates for virtual element methods in three dimensions, Comput. Methods Appl. Math. 23 (1) (2023) 177–187.

[35] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, third ed., in: Texts in Applied Mathematics, vol. 15, Springer, New York, 2008.

[36] Y. Wang, S. Rhebergen, Space–time hybridizable discontinuous Galerkin method for advection-diffusion on deforming domains: The advection-dominated regime, 2023, arXiv:2308.12130.

[37] L. Diening, J. Storn, T. Tscherpel, Interpolation operator on negative Sobolev spaces, Math. Comp. 92 (342) (2023) 1511–1541.

[38] A. Cangiani, E.H. Georgoulis, T. Pryer, O.J. Sutton, A posteriori error estimates for the virtual element method, Numer. Math. 137 (4) (2017) 857–893.

[39] L. Chen, J. Huang, Some error analysis on virtual element methods, Calcolo 55 (1) (2018) 23, Paper No. 5.

[40] P. Grisvard, Elliptic Problems in Nonsmooth Domains, in: Classics in Applied Mathematics, vol. 69, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[41] L. Beirão da Veiga, F. Dassi, S. Gómez, SUPG-stabilized time-DG finite and virtual elements for the time-dependent advection-diffusion equation, 2024, arXiv:2410.13635.

[42] A. Henderson, ParaView Guide, a Parallel Visualization Application, Kitware Inc., 2007.